

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2003, 17:51:00 ; Search time 404 Seconds
(without alignments)
8501.222 Million cell updates/sec

Title: US-09-872-523-6

Perfect score: 1497
Sequence: 1 atgtctgaattctgtaaat.....cgaaaagracacattga 1497

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1533700 seqs, 1147125425 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3067400

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published Applications_NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1497	100.0	1497	10	US-09-872-523-6
2	1446	96.6	1497	10	US-09-872-523-73
3	1446	96.6	1497	10	US-09-872-523-74
4	1446	96.6	1497	10	US-09-872-523-75
5	833	55.6	2307	10	US-09-872-523-76
6	833	55.6	2307	10	US-09-872-523-77
7	833	55.6	2307	10	US-09-872-523-78
8	20	1.3	280	9	US-09-864-761-31987
9	20	1.3	425	9	US-09-864-761-15473
10	20	1.3	2863	13	US-10-136-224-2
11	20	1.3	3686	13	US-10-136-224-3
12	20	1.3	3940	13	US-10-136-224-1
13	20	1.3	11996	12	US-10-240-485-46
14	19	1.3	7450	12	US-10-240-453-6
15	19	1.3	27190	14	US-10-195-144-87
16	19	1.3	322101	12	US-10-060-902-1

17	18	1.2	286	10	US-09-989-920-127	Sequence 127, App
18	18	1.2	388	11	US-09-918-995-37801	Sequence 37801, A
19	18	1.2	409	9	US-09-864-761-16621	Sequence 16621, A
c	20	1.2	451	13	US-10-027-632-2047	Sequence 2047, Ap
21	18	1.2	491	10	US-09-796-692-7266	Sequence 7266, Ap
22	18	1.2	491	14	US-10-040-862-7266	Sequence 7266, Ap
23	18	1.2	559	13	US-10-027-633-257290	Sequence 257290, Ap
24	18	1.2	611	13	US-10-027-633-225568	Sequence 225568, Ap
25	18	1.2	611	13	US-10-027-633-225569	Sequence 225569, Ap
26	18	1.2	611	13	US-10-027-633-225570	Sequence 225570, Ap
c	27	1.2	639	14	US-10-198-846-8267	Sequence 8267, Ap
28	18	1.2	640	13	US-10-027-633-192293	Sequence 192293, Ap
29	18	1.2	640	13	US-10-027-633-192294	Sequence 192294, Ap
30	18	1.2	650	13	US-10-027-633-228137	Sequence 228137, Ap
31	18	1.2	650	13	US-10-027-633-228138	Sequence 228138, Ap
c	32	1.2	711	10	US-09-816-828-11	Sequence 11, Appl
33	18	1.2	742	11	US-09-984-271-76	Sequence 76, Appl
c	34	1.2	781	13	US-10-202-193-321	Sequence 321, Appl
35	18	1.2	906	14	US-10-198-846-13849	Sequence 13849, A
36	18	1.2	981	13	US-10-027-633-153545	Sequence 153545, Ap
37	18	1.2	1014	13	US-10-027-633-121892	Sequence 121892, Ap
38	18	1.2	1172	9	US-09-939-980-8	Sequence 8, Appl
39	18	1.2	1620	14	US-10-128-714-2459	Sequence 2459, Ap
40	18	1.2	1671	14	US-10-128-714-1459	Sequence 1459, Ap
41	18	1.2	1707	14	US-10-128-714-7459	Sequence 7459, Ap
42	18	1.2	1812	14	US-10-128-714-6459	Sequence 6459, Ap
43	18	1.2	3671	14	US-10-128-714-459	Sequence 459, Ap
c	44	1.2	3767	14	US-10-103-313-270	Sequence 270, Appl
45	18	1.2	3812	14	US-10-128-714-3459	Sequence 3459, Ap

ALIGNMENTS

RESULT 1	
US-09-872-523-6	
Sequence 6, Application US/09872523	
Patent No. US20020137906A1	
GENERAL INFORMATION:	
APPLICANT: Davison, H. Robert	
APPLICANT: Davison, Eva M.	
APPLICANT: Lu, Xiaowei	
TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans	
FILE REFERENCE: 01997/536002	
CURRENT APPLICATION NUMBER: US/09/872,523	
CURRENT FILING DATE: 2001-06-01	
PRIOR APPLICATION NUMBER: US 60/208,802	
PRIOR FILING DATE: 2000-06-02	
NUMBER OF SEQ ID NOS: 78	
SOFTWARE: FastSeq for Windows Version 4.0	
SEQ ID NO 6	
LENGTH: 1497	
TYPE: DNA	
ORGANISM: Caenorhabditis elegans	
US-09-872-523-6	
Query Match	100.0% Score 1497; DB 10; Length 1497;
Best Local Similarity	100.0% Pred. No. 0;
Matches 1497; Conservative	0; Mismatches
	0; Indels
	0; Gaps
	0;
Qy	1 ATGCTGAATTTCTGAAAATTTGTCAGAGCTAACAAAAAATCGACAGAAAACTCGATTAAG 60
Db	1 ATGCTGAATTTCTGAAAATTTGTCAGAGCTAACAAAAAATCGACAGAAAACTCGATTAAG 60
Qy	61 ACCTACTTGGGAAATCCTTTTACATCACTTGGAGAAAGAAAACTTTTCATTCCA 120
Db	61 ACCTACTTGGGAAATCCTTTTACATCACTTGGAGAAAGAAAACTTTTCATTCCA 120
Qy	121 GTTGAAGCATTCATCGTACCTTACAGTTAATTTTAACGATGCGTGAAGAGGATT 180
Db	121 GTTGAAGCATTCATCGTACCTTACAGTTAATTTTAACGATGCGTGAAGAGGAGTT 180
Qy	181 ATCTCGAAGCAAGTGCTCATGATTATGACAAAGCACTGCATTCAAGTCAGATGG 240

|||||
Db 181 ATCTCGAAACAGGCGCCATGATTATGACAAAGAACTCGATTCGATTCAAGTACAGATGG 240
OY 241 TTTTCAGCAATGTAAGAAAGTTGGGATACAGAGTTCGGCTCAGTTTATGGAGCTGAC 300
Db 241 TTTTCAGCAATGTAAGAAAGTTGGGATACAGAGTTCGGCTCAGTTTATGGAGCTGAC 300
OY 301 ACGAAATTTTGGCTCAATATTTTATCGACGATATGTTTGGTTTGGCAAGCGCCGAATG 360
Db 301 ACGAAATTTTGGCTCAATATTTTATCGACGATATGTTTGGTTTGGCAAGCGCCGAATG 360
OY 361 AGTATGCCAATATGATGTAATAATTTGATATCTCCGCGCTTGCATATACGAAGAATAC 420
Db 361 AGTATGCCAATATGATGTAATAATTTGATATCTCCGCGCTTGCATATACGAAGAATAC 420
OY 421 CAAATGATATGTAATATTTATGTAATTTGATATGATGCGGAATGCGGCCAAACT 480
Db 421 CAAATGATATGTAATATTTATGTAATTTGATATGATGCGGAATGCGGCCAAACT 480
OY 481 TCGCTGTCTCCAAAATTCGATGAAGGAGGCTCTCCTTACGCAAGCATCGTTTCAAAGTT 540
Db 481 TCGCTGTCTCCAAAATTCGATGAAGGAGGCTCTCCTTACGCAAGCATCGTTTCAAAGTT 540
OY 541 GGACACGCTTTGACATATTTAATTTATCTCAATTTCTACTGAATATCGGCTAGCGGAT 600
Db 541 GGACACGCTTTGACATATTTAATTTATCTCAATTTCTACTGAATATCGGCTAGCGGAT 600
OY 601 CAAGAATATGTGACGACGACGATGATCTATCAACAAGAAAGATTTCCGCAATCG 660
Db 601 CAAGAATATGTGACGACGACGATGATCTATCAACAAGAAAGATTTCCGCAATCG 660
OY 661 CTTCAGATGACATGACGACGACAGTCTTACCTGCTGATCTCAATATTTGATAGAC 720
Db 661 CTTCAGATGACATGACGACGACAGTCTTACCTGCTGATCTCAATATTTGATAGAC 720
OY 721 GAGGGAACCTTCTCATATTTCTCGTTGATTTGAGAGTTCAGCAATGATCAACTAAT 780
Db 721 GAGGGAACCTTCTCATATTTCTCGTTGATTTGAGAGTTCAGCAATGATCAACTAAT 780
OY 781 GCGAAAAAGCAATATATTTGACACACAATAATTTGCTCAACGAATAAAAATGGAGAA 840
Db 781 GCGAAAAAGCAATATATTTGACACACAATAATTTGCTCAACGAATAAAAATGGAGAA 840
OY 841 AATCCAAATATGACTCGACGACGCTCATTGATCAATTAAGCAAAAAGATCCAAATGAT 900
Db 841 AATCCAAATATGACTCGACGACGCTCATTGATCAATTAAGCAAAAAGATCCAAATGAT 900
OY 901 CCCATGATTTGGAGAAAGTTAAGGTTGAGCAAAAAGTTGAGCTGATGACCCCTTGGCT 960
Db 901 CCCATGATTTGGAGAAAGTTAAGGTTGAGCAAAAAGTTGAGCTGATGACCCCTTGGCT 960
OY 961 CAGCAATTCATTAACCTCCACGCTCGCTTCGATTTCTCAAAATTTTGGCAAAAGTGAAGATAT 1020
Db 961 CAGCAATTCATTAACCTCCACGCTCGCTTCGATTTCTCAAAATTTTGGCAAAAGTGAAGATAT 1020
OY 1021 CTTATATGGGAAATGATGATGCTCAGATGCTGTAAGACAGTTTCCATTCATATCAAT 1080
Db 1021 CTTATATGGGAAATGATGATGCTCAGATGCTGTAAGACAGTTTCCATTCATATCAAT 1080
OY 1081 AATACATTTATGTTCCAGTTGGTTATGCGGAAAGTAAATTTGGAAGCTGTTCCGCCA 1140
Db 1081 AATACATTTATGTTCCAGTTGGTTATGCGGAAAGTAAATTTGGAAGCTGTTCCGCCA 1140
OY 1141 GATGAGTTCAAGAAACATTCAGATGATGATCTTGCAGAAAGAAATTCGCAAGAAC 1200
Db 1141 GATGAGTTCAAGAAACATTCAGATGATGATCTTGCAGAAAGAAATTCGCAAGAAC 1200
OY 1201 CTACCGCTTGACTTGTTCAGCAATGCTTCCCAAGAGAGATTAGCAAAATTTAAGGTA 1260
Db 1201 CTACCGCTTGACTTGTTCAGCAATGCTTCCCAAGAGAGATTAGCAAAATTTAAGGTA 1260
OY 1261 ATTCTGATTTCCAAACGGGTAGAGCTAGCCCTTGAAGCTGCTGACATGTGAAAAATCAG 1320
Db 1261 ATTCTGATTTCCAAACGGGTAGAGCTAGCCCTTGAAGCTGCTGACATGTGAAAAATCAG 1320

Db 1261 ATTCTGATTTCCAAACGGGTAGAGCTAGCCCTGGAAGCTGCTGACATGTGAAAAATCAG 1320
OY 1321 TTTATTTTGTCCAGCTACAGTGAATAATCAGTTTCATGGAAGACTGATTAATGTCATTTGAC 1380
Db 1321 TTTATTTTGTCCAGCTACAGTGAATAATCAGTTTCATGGAAGACTGATTAATGTCATTTGAC 1380
OY 1381 GCGTGGATGAAAGAAATTTGATGAACCTGATGATGTGACACTCCATGATTTCTACGATA 1440
Db 1381 GCGTGGATGAAAGAAATTTGATGAACCTGATGATGTGACACTCCATGATTTCTACGATA 1440
OY 1441 GGATGGTGAAGGACAGATTTATGTTCTACACCTCCGAAAAAGTAACTATTTGA 1497
Db 1441 GGATGGTGAAGGACAGATTTATGTTCTACACCTCCGAAAAAGTAACTATTTGA 1497
RESULT 2
US-09-872-523-73
; Sequence 73, Application US/09872523
; Patent No. US20020137906A1
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Davison, Ewa M.
; APPLICANT: Lu, Xiaowei
; TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
; FILE REFERENCE: 01997/536002
; CURRENT APPLICATION NUMBER: US/09/872, 523
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/208, 802
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 1497
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-872-523-73
Query Match 96.6%; Score 1446; DB 10; Length 1497;
Best Local Similarity 99.9%; Freq. No. 0;
Matches 1496; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 ATGCTGAATTTTGGAAATTTGTCAGAGCTTAACAAAAATGCGACAGAAACTGATTAAG 60
Db 1 ATGCTGAATTTTGGAAATTTGTCAGAGCTTAACAAAAATGCGACAGAAACTGATTAAG 60
OY 61 ACCTACTGTGGGAATCCTATTTACATCAGTTTCGAGAAAGAAACTTCTTCAATCCA 120
Db 61 ACCTACTGTGGGAATCCTATTTACATCAGTTTCGAGAAAGAAACTTCTTCAATCCA 120
OY 121 GTTGAAGATTCATTCGTAACCTTAAGTTTAATTTAAGCAATGCGTGAAGAGGAGTT 180
Db 121 GTTGAAGATTCATTCGTAACCTTAAGTTTAATTTAAGCAATGCGTGAAGAGGAGTT 180
OY 181 ATCTTCGAACAGTGTGCTCATGATTTATGACAAGAACTCGATTTGATTTCAAGTCAATGG 240
Db 181 ATCTTCGAACAGTGTGCTCATGATTTATGACAAGAACTCGATTTGATTTCAAGTCAATGG 240
OY 241 TTTGACGCAATTTGAAAAAGTTTGGCGATACAGAGTCTGCTCAGATTATTCGAGCTGAC 300
Db 241 TTTGACGCAATTTGAAAAAGTTTGGCGATACAGAGTCTGCTCAGATTATTCGAGCTGAC 300
OY 301 ACGAAATTTTGGCTCAATATTTTATCGGACGATATGTTTGGTTTGGCAAGCGCCGAATG 360
Db 301 ACGAAATTTTGGCTCAATATTTTATCGGACGATATGTTTGGTTTGGCAAGCGCCGAATG 360
OY 361 AGTATGCCAATATGATGTAATAATTTGATATCTCCGCGCTTGCATATACGAAGAATAC 420
Db 361 AGTATGCCAATATGATGTAATAATTTGATATCTCCGCGCTTGCATATACGAAGAATAC 420
OY 421 CAAATGATATGTAATATTTATGTAATTTGATATGATGCGGAATGCGGCCAAACT 480
Db 421 CAAATGATATGTAATATTTATGTAATTTGATATGATGCGGAATGCGGCCAAACT 480

```

OY 481 TCGCTGTCCTCAAAATTCGATGAAGGAGGCTCTCTCTTAAGCAAGCATGCTTTCAAAGTT 540
    |||||||
Db 481 TCGCTGTCCTCAAAATTCGATGAAGGAGGCTCTCTCTTAAGCAAGCATGCTTTCAAAGTT 540
OY 541 GGACAAACGCTTGAACTATTAAATTATTTCCATTCTACTGAAATACGGTAGCGCAATT 600
    |||||||
Db 541 GGACAAACGCTTGAACTATTAAATTATTTCCATTCTACTGAAATACGGTAGCGCAATT 600
OY 601 CAAGAAATATGTGACAGCAAGATGATGATCATACAAAGAAAGCTTCCCGCAATG 660
    |||||||
Db 601 CAAGAAATATGTGACAGCAAGATGATGATCATACAAAGAAAGCTTCCCGCAATG 660
OY 661 CTTCAGATGATGATGACAGCAAGCAAGCTTTAGCTGATCTCAATATTTGATGATAGC 720
    |||||||
Db 661 CTTCAGATGATGATGACAGCAAGCAAGCTTTAGCTGATCTCAATATTTGATGATAGC 720
OY 721 GAGGGAAGCTTCTTCATATTTCTGTTGATTTGACGAGCTCAATGATATCACTAAT 780
    |||||||
Db 721 GAGGGAAGCTTCTTCATATTTCTGTTGATTTGACGAGCTCAATGATATCACTAAT 780
OY 781 GCGAAAAAGCAATATATTTAGACACACAAATTAATTTGCTCAAGCAATTAATAATGAGAA 840
    |||||||
Db 781 GCGAAAAAGCAATATATTTAGACACACAAATTAATTTGCTCAAGCAATTAATAATGAGAA 840
OY 841 AATCCAAAGATATGACTCAGACAGCAAGCTCAGATTTGATCAATTAGCAAAAGATCCAAATGAT 900
    |||||||
Db 841 AATCCAAAGATATGACTCAGACAGCAAGCTCAGATTTGATCAATTAGCAAAAGATCCAAATGAT 900
OY 901 CCCATGATTTGGAGAAAAATTAAGTTGGACCAAAAATTTGAGCTCTGCAACCCCTTGCT 960
    |||||||
Db 901 CCCATGATTTGGAGAAAAATTAAGTTGGACCAAAAATTTGAGCTCTGCAACCCCTTGCT 960
OY 961 CAGCAATTCGAATTAACCTCAGCTCGCTGATTTCTCAATTTTGCAAAATGAGAGATAT 1020
    |||||||
Db 961 CAGCAATTCGAATTAACCTCAGCTCGCTGATTTCTCAATTTTGCAAAATGAGAGATAT 1020
OY 1021 CTTATTTGGGAAATGAGATGCTCCAGATGCACTTGAAGACAGTTTCTATTCATCAAT 1080
    |||||||
Db 1021 CTTATTTGGGAAATGAGATGCTCCAGATGCACTTGAAGACAGTTTCTATTCATCAAT 1080
OY 1081 AATPACTTATGTCTCCAGTGTGTTATGGGAAAAATTAATTTGAACTTGTTCGCGCA 1140
    |||||||
Db 1081 AATPACTTATGTCTCCAGTGTGTTATGGGAAAAATTAATTTGAACTTGTTCGCGCA 1140
OY 1141 GATGAGTTCAAAAGCAATTCAGATGAGATGATGATCTTGAGAAAAATCTCGAGAAACC 1200
    |||||||
Db 1141 GATGAGTTCAAAAGCAATTCAGATGAGATGATGATCTTGAGAAAAATCTCGAGAAACC 1200
OY 1201 CTACCGCTTGACTTGTTCAGGCCAATGCTTCCCAAGAGAGATTAGCAAAATTTAAGTA 1260
    |||||||
Db 1201 CTACCGCTTGACTTGTTCAGGCCAATGCTTCCCAAGAGAGATTAGCAAAATTTAAGTA 1260
OY 1261 AATTCGATTTCCAAAGGGGTAGGACTACGCCCTTGAAGCTGCTGACATGTGTGAAATCAG 1320
    |||||||
Db 1261 AATTCGATTTCCAAAGGGGTAGGACTACGCCCTTGAAGCTGCTGACATGTGTGAAATCAG 1320
OY 1321 TTTATTTGTCACAGTACAGTGAATCAGTTTCATGGAAGACTGTAATATGCAATTTGCGAC 1380
    |||||||
Db 1321 TTTATTTGTCACAGTACAGTGAATCAGTTTCATGGAAGACTGTAATATGCAATTTGCGAC 1380
OY 1381 GGCCTGGAGTGAAGAAATTTGATGAACGTGATGATGTGACCTCCATGATATTCACGATA 1440
    |||||||
Db 1381 GGCCTGGAGTGAAGAAATTTGATGAACGTGATGATGTGACCTCCATGATATTCACGATA 1440
OY 1441 GGAATGCTGAAGCGACAGCTATATGTTCTACAAACCTCGCAAAAAGTAACTAATGA 1497
    |||||||
Db 1441 GGAATGCTGAAGCGACAGCTATATGTTCTACAAACCTCGCAAAAAGTAACTAATGA 1497

```

RESULT 3

US-09-872-523-74

; Sequence 74, Application US/09872523
; Patent No. US20020137906A1

```

; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Davison, Ewa M.
; APPLICANT: Lu, Xiaowei
; TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
; FILE REFERENCE: 01997/536002
; CURRENT APPLICATION NUMBER: US/09/872,523
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/208,802
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 1497
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-872-523-74

Query Match          96.6%; Score 1446; DB 10; Length 1497;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1496; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ATGCTGAATTTCTGAAAATTTGTCAGAGCTTAACAAAAATCGGACGAAAATCTCGATAAG 60
    |||||||
Db 1 ATGCTGAATTTCTGAAAATTTGTCAGAGCTTAACAAAAATCGGACGAAAATCTCGATAAG 60
OY 61 ACCTACTTGGGAAATCTATTTACATGATTCAGATTCGAGAAAGAAAATCTTTCATTCCA 120
    |||||||
Db 61 ACCTACTTGGGAAATCTATTTACATGATTCAGATTCGAGAAAGAAAATCTTTCATTCCA 120
OY 121 GTTGAAGCATTCGAATGTAACCTTACAGTTAATTTTAAAGATGCTGGAAGAGAGTT 180
    |||||||
Db 121 GTTGAAGCATTCGAATGTAACCTTACAGTTAATTTTAAAGATGCTGGAAGAGAGTT 180
OY 181 ACTTCGAAACAGTGTGTCATGATTAATGACAAAGAACTGGATTCGATTAAGCAGATGG 240
    |||||||
Db 181 ACTTCGAAACAGTGTGTCATGATTAATGACAAAGAACTGGATTCGATTAAGCAGATGG 240
OY 241 TTTGCAAGATTTGAAAAGTTTGCAGATACAGAGTTCTGGCTCAGTTATCGAGCTGAC 300
    |||||||
Db 241 TTTGCAAGATTTGAAAAGTTTGCAGATACAGAGTTCTGGCTCAGTTATCGAGCTGAC 300
OY 301 ACGAAATTTGGCTCAATATTTTATCGACGATATGTTGGTTGCAAAACGCCGCAATG 360
    |||||||
Db 301 ACGAAATTTGGCTCAATATTTTATCGACGATATGTTGGTTGCAAAACGCCGCAATG 360
OY 361 AAGTATCCCAATATGATTAATTTGATATGCTCCGCCCTTGCAATCAACGAAGATAC 420
    |||||||
Db 361 AAGTATCCCAATATGATTAATTTGATATGCTCCGCCCTTGCAATCAACGAAGATAC 420
OY 421 CAAATGATATGTAATTTATGTAATTAATTTGATGATGAGCGCAATTCGCGCAAACT 480
    |||||||
Db 421 CAAATGATATGTAATTTATGTAATTAATTTGATGATGAGCGCAATTCGCGCAAACT 480
OY 481 TCGCTGTCCTCAAAATTCGATGAAGGAGGCTCTCTTAAGCAAGCATGCTTTCAAAGTT 540
    |||||||
Db 481 TCGCTGTCCTCAAAATTCGATGAAGGAGGCTCTCTTAAGCAAGCATGCTTTCAAAGTT 540
OY 541 GGACAAACGCTTGAACTATTAAATTATTTCCATTCTACTGAAATACGGTAGCGCAATT 600
    |||||||
Db 541 GGACAAACGCTTGAACTATTAAATTATTTCCATTCTACTGAAATACGGTAGCGCAAAAT 600
OY 601 CAAGAAATATGTGACAGCAAGATGATGATCATACAAAGAAAGCTTCCCGCAATG 660
    |||||||
Db 601 CAAGAAATATGTGACAGCAAGATGATGATCATACAAAGAAAGCTTCCCGCAATG 660
OY 661 CTTCAGATGATGATGACAGCAAGCAAGCTTTAGCTGATCTCAATATTTGATGATAGC 720
    |||||||
Db 661 CTTCAGATGATGATGACAGCAAGCAAGCTTTAGCTGATCTCAATATTTGATGATAGC 720
OY 721 GAGGGAAGCTTCTTCATATTTCTGTTGATTTGACGAGCTCAATGATATCACTAAT 780
    |||||||
Db 721 GAGGGAAGCTTCTTCATATTTCTGTTGATTTGACGAGCTCAATGATATCACTAAT 780

```


Db	1021	CTTATTGTGGGAATGANGTGGCCAGATGCACTTTGAAGACAGTTTCTCATTCATATCAAT	1080
QY	1081	AATACATTTATGTTCCAGTGTGTTATGCGGAAAAGTATAATTTTGAACCTGTTCCGCCA	1140
Db	1081	AATACATTTATGTTCCAGTGTGTTATGCGGAAAAGTATAATTTTGAACCTGTTCCGCCA	1140
QY	1141	GATAGTTCAAAGGAACATTCAGATGGATGGATGAATACCTTGGAGAAAAGAATCTGCAGAAACC	1200
Db	1141	GATAGTTCAAAGGAACATTCAGATGGATGGATGAATACCTTGGAGAAAAGAATCTGCAGAAACC	1200
QY	1201	CTACCGCTTGACCTGTTTCAAGCCCAATGCGTTCCTCCAGAGAGATTTAGCAAAATTTAAGSTA	1260
Db	1201	CTACCGCTTGACCTGTTTCAAGCCCAATGCGTTCCTCCAGAGAGATTTAGCAAAATTTAAGSTA	1260
QY	1261	ATTCTGATTTTCCAAACGGGATGAGCACTACGCCCTTGAAGCTGCTGCATGTGTGAATAATCAG	1320
Db	1261	ATTCTGATTTTCCAAACGGGATGAGCACTACGCCCTTGAAGCTGCTGCATGTGTGAATAATCAG	1320
QY	1321	TTTATTGTTGTCACGCTACAGTGAATTCAGTTTCATGGAAGACATGATTAATTCATTTTGAC	1380
Db	1321	TTTATTGTTGTCACGCTACAGTGAATTCAGTTTCATGGAAGACATGATTAATTCATTTTGAC	1380
QY	1381	GGCTGGGATGAAGAATTTGATGAACCTGTATGATGTGGACCTCCCATGATATTCTACCGGATA	1440
Db	1381	GGCTGGGATGAAGAATTTGATGAACCTGTATGATGTGGACCTCCCATGATATTCTACCGGATA	1440
QY	1441	GGATGGGTGTGAAGCGCACAGTTATGTTCTTCAACACCTCCGAAAAAGTACAACTATTGGA	1497
Db	1441	GGATGGGTGTGAAGCGCACAGTTATGTTCTTCAACACCTCCGAAAAAGTACAACTATTGGA	1497

```

RESULT 5
US-09-872-523-76
; Sequence 76, Application US/09872523
; Patent No. US20020137906A1
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Davison, Ewa M.
; APPLICANT: Lu, Xiaowei
; TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
; FILE REFERENCE: 01997/536002
; CURRENT APPLICATION NUMBER: US/09/872,523
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/208,802
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 2307
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; US-09-872-523-76

```

Query Match	55.6%	Score 833;	DB 10;	Length 2307;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 833;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	448	AATTGACATGTGATGGCGAAATCGTCGGCGAAACCTTCGCGTCTCCAAATTCGATGAAGG	507
Db	1138	AATTGACATGTGATGGCGAAATCGTCGGCGAAACCTTCGCGTCTCCAAATTCGATGAAGG	1197
QY	508	AAGGCTCTCCGTAAGCAAGCATCGTTTCCAAAGTTGGACAAAGCTCTTGAACATTAATAATAT	567
Db	1198	AAGGCTCTCCGTAAGCAAGCATCGTTTCCAAAGTTGGACAAAGCTCTTGAACATTAATAATAT	1255
QY	568	TCCAAATTCACGTAATACGCGTACGCGGAAATTCAGAATAATATGTGGACGACGATGAAT	627
Db	1258	TCCAAATTCACGTAATACGCGTACGCGGAAATTCAGAATAATATGTGGACGACGATGAAT	1317
QY	628	GTATCTATCACAAGAAAGAACTTCCCGGAATCGCTTCCAGATGCGAGATGACGACAGACAA	687
Db	1318	GTATCTATCACAAGAAAGAACTTCCCGGAATCGCTTCCAGATGCGAGATGACGACAGACAA	1377

QY	688	GTCTTTGACGTCGATGCTCATATTGGATATACGAGGAAGCTTCTTCATATTCCGTGT	747
Db	1378	GTCTTTAGCTGTGGATCTCAATATTGGATACGAGGGAACCTTCTCATATTCCGTGT	1437
OY	748	GGATTTCGACGACGATCGATGATTCACATTAATCGAAAAAGAAATATTATTGAGCACACA	807
Db	1438	GGATTTCGACGACGATCGATGATTCACATTAATCGAAAAAGAAATATTATTGAGCACACA	1497
OY	808	AATAAAAATTGCTCAAGCAATAAAAAATGAGAAAAATCCAAAGATATGACTCAGACAGCTC	867
Db	1498	AATAAAAATTGCTCAAGCAATAAAAAATGAGAAAAATCCAAAGATATGACTCAGACAGCTC	1557
OY	868	ACATTGTGATCATATTAGCAAAAAGATCCCAATTGATCCCATGATTTGGAGAAAAAGTTAAGGTT	927
Db	1558	ACATTGTGATCATATTAGCAAAAAGATCCCAATTGATCCCATGATTTGGAGAAAAAGTTAAGGTT	1617
OY	928	GGACAAAAAGTTTGAGCTCATCGACCCCTTGCGCTCAGAAATTCATTAACCTCCAGCTGCCT	987
Db	1618	GGACAAAAAGTTTGAGCTCATCGACCCCTTGCGCTCAGCAATTCATTAACCTCCAGCTGCCT	1677
OY	988	TCGATTTCTCAAAATTTTGCAAAACGTGAAGAGATATCTTATTGTGGGAATGATGTGCCAGAT	1047
Db	1678	TCGATTTCTCAAAATTTTGCAAAACGTGAAGAGATATCTTATTGTGGGAATGATGTGCCAGAT	1737
OY	1048	GCACCTTGAAGCACTTTTCTCATTCATATCAATTAATACATTTAATGTTCCCGAGTGGTAT	1107
Db	1738	GCACCTTGAAGCACTTTTCTCATTCATATCAATTAATACATTTAATGTTCCCGAGTGGTAT	1797
OY	1108	GGGAAAAAGTTAATTTTGGAACCTGTTCCCGCAGATGAGTTCAAGAACAATTCAGATGG	1167
Db	1798	GGGAAAAAGTTAATTTTGGAACCTGTTCCCGCAGATGAGTTCAAGAACAATTCAGATGG	1857
OY	1168	GATGAATACCTTGAGAAAAAGAACTGCAAAAACCTACCGCTTGACTTGTTCAAAGCCAAATG	1227
Db	1858	GATGAATACCTTGAGAAAAAGAACTGCAAAAACCTACCGCTTGACTTGTTCAAAGCCAAATG	1917
OY	1228	CCTTCCCAAGAGAAATTTAGACAATTTTAAGTATTTCGATTTCCAAAGGGGT	1280
Db	1918	CCTTCCCAAGAGAAATTTAGACAATTTTAAGTATTTCGATTTCCAAAGGGGT	1970

```

RESULT 6
US-09-872-523-77
; Sequence 77, Application US/09872523
; Patent No. US20020137906A1
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Davison, Ewa M.
; APPLICANT: Lu, Xiaowei
; TITLE OF INVENTION: A Tumor Suppressor Pathway In C. Elegans
; FILE REFERENCE: 01997/536002
; CURRENT APPLICATION NUMBER: US/09/872.523
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/208,802
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77
; LENGTH: 2307
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; US-09-872-523-77

```

	Query Match	55.6%	Score 833	DB 10	Length 2307	
	Best Local Similarity	100.0%	Pred. NO. 0			
	Matches 833	Conservative 0	Mismatches 0	Indels 0	Gaps 0	
QY	448	AATTGCAATTGATGCGAAGATGTCGGCGCAAACTTCGCTGCTCCAAAATTGCATGAAGG	507			
Db	1138	AATTGCATGTGATGCGAAGATGTCGGCGCAAACTTCGCTGCCAAAATTGCATGAAGG	1197			
QY	508	AAGGCTTCTCAAGCAAGCATGTTTCAAAAGTTGACACAGCTTTGAACATTTAAATAT	567			

```
Db 1198 AAGGCTCTCTAAGCAAGCATCTTTCAAAAGTTGACAAAGCTTGTGAACATATTAATAT 1257
Oy 568 TCCAAATTTCTACTGAATAACGCGTAGCGCAATTTCAAGAAATATGTGACGAGCAATGAT 627
Db 1258 TCCAAATTTCTACTGAATAACGCGTAGCGCAATTTCAAGAAATATGTGACGAGCAATGAT 1317
Oy 628 GTATCTATCACAAAGAAAGACTTTTCCCGATGCTTCCAGATGACAGATGACAGACACA 687
Db 1318 GTATCTATCACAAAGAAAGACTTTTCCCGATGCTTCCAGATGACAGATGACAGACACA 1377
Oy 688 GTCTTTAGCTCTGATCTCAATATTTGAGATGACGAGGAGGAGCTTCTTCATATTTCTGTT 747
Db 1378 GTCTTTAGCTCTGATCTCAATATTTGAGATGACGAGGAGGAGCTTCTTCATATTTCTGTT 1437
Oy 748 GGATTTGACAGCATCAATGATATCACTAAATGCGAAAGAAAGAAATATTTGAGACACA 807
Db 1438 GGATTTGACAGCATCAATGATATCACTAAATGCGAAAGAAAGAAATATTTGAGACACA 1497
Oy 808 AATTAATATGCTCAAGCATATAAAAAATGAGAAAAATCCAAAGATATGACTCAGACAGCGTC 867
Db 1498 AATTAATATGCTCAAGCATATAAAAAATGAGAAAAATCCAAAGATATGACTCAGACAGCGTC 1557
Oy 868 ACATTTGATCAATTTAGCAAAAGATCCAAATTTGATCCCATGATTTGGAGAAAGTTAAGTT 927
Db 1558 ACATTTGATCAATTTAGCAAAAGATCCAAATTTGATCCCATGATTTGGAGAAAGTTAAGTT 1617
Oy 928 GGACAAAGTTTGAAGCTACGACCCCTTGCTCAGCAATTCATTAACCTCCAGCTGCT 987
Db 1618 GGACAAAGTTTGAAGCTACGACCCCTTGCTCAGCAATTCATTAACCTCCAGCTGCT 1677
Oy 988 TCGATTTCAAAATTTTGGCAAAAGTGAAGATATCTTATTTGGGAATGGATGATGCTCAGAT 1047
Db 1678 TCGATTTCAAAATTTTGGCAAAAGTGAAGATATCTTATTTGGGAATGGATGATGCTCAGAT 1737
Oy 1048 GCACCTTGAAGACAGTTTCTATTCATATCAATTAATACATTTATGTTCCCAAGTTGTTAT 1107
Db 1738 GCACCTTGAAGACAGTTTCTATTCATATCAATTAATACATTTATGTTCCCAAGTTGTTAT 1797
Oy 1108 GCGGAAAAGTATATTTGGAACTTGTCCGCCAGATGAGTTCAAGAAACATTCAGATGG 1167
Db 1798 GCGGAAAAGTATATTTGGAACTTGTCCGCCAGATGAGTTCAAGAAACATTCAGATGG 1857
Oy 1168 GATGAATACTTGGAGAAAGATCTGCGAAGACCTTACCGCTTGACTTGTTCAGACCAATG 1227
Db 1858 GATGAATACTTGGAGAAAGATCTGCGAAGACCTTACCGCTTGACTTGTTCAGACCAATG 1917
Oy 1228 CCTTCCCAAGAGAGATTAGCAAAATTTAAGTAATTTCTGATTTCCAAACGGGT 1280
Db 1918 CCTTCCCAAGAGAGATTAGCAAAATTTAAGTAATTTCTGATTTCCAAACGGGT 1970

RESULT 7
US-09-872-523-78
; Sequence 78, Application US/09872523
; Patent No. US20020137906A1
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Davison, Eva M.
; APPLICANT: Lu, Xiaowei
; TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
; FILE REFERENCE: 01997/536002
; CURRENT APPLICATION NUMBER: US/09/872,523
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/208,802
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 2307
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-872-523-78
```

```
Query Match 55.6%; Score 833; DB 10; Length 2307;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 448 AATGCAATGATGCGAAATGTCGGCCAACTTCGCTGCTCCAAATTTGATGAAGG 507
Db 1138 AATGCAATGATGCGAAATGTCGGCCAACTTCGCTGCTCCAAATTTGATGAAGG 1197
Oy 508 AAGGCTCTCTAAGCAAGCATCTTTCAAAAGTTGACAAAGCTTGTGAACATATTAATAT 567
Db 1198 AAGGCTCTCTAAGCAAGCATCTTTCAAAAGTTGACAAAGCTTGTGAACATATTAATAT 1257
Oy 568 TCCAAATTTCTACTGAATAACGCGTAGCGCAATTTCAAGAAATATGTGACGAGCAATGAT 627
Db 1258 TCCAAATTTCTACTGAATAACGCGTAGCGCAATTTCAAGAAATATGTGACGAGCAATGAT 1317
Oy 628 GTATCTATCACAAAGAAAGACTTTTCCCGATGCTTCCAGATGACAGATGACAGACACA 687
Db 1318 GTATCTATCACAAAGAAAGACTTTTCCCGATGCTTCCAGATGACAGATGACAGACACA 1377
Oy 688 GTCTTTAGCTCTGATCTCAATATTTGATGACGAGGAGGAGCTTCTTCATATTTCTGTT 747
Db 1378 GTCTTTAGCTCTGATCTCAATATTTGATGACGAGGAGGAGCTTCTTCATATTTCTGTT 1437
Oy 748 GGATTTGACAGCATCAATGATATCACTAAATGCGAAAGAAAGAAATATTTGAGACACA 807
Db 1438 GGATTTGACAGCATCAATGATATCACTAAATGCGAAAGAAAGAAATATTTGAGACACA 1497
Oy 808 AATTAATATGCTCAAGCATATAAAAAATGAGAAAAATCCAAAGATATGACTCAGACAGCGTC 867
Db 1498 AATTAATATGCTCAAGCATATAAAAAATGAGAAAAATCCAAAGATATGACTCAGACAGCGTC 1557
Oy 868 ACATTTGATCAATTTAGCAAAAGATCCAAATTTGATCCCATGATTTGGAGAAAGTTAAGTT 927
Db 1558 ACATTTGATCAATTTAGCAAAAGATCCAAATTTGATCCCATGATTTGGAGAAAGTTAAGTT 1617
Oy 928 GGACAAAGTTTGAAGCTACGACCCCTTGCTCAGCAATTCATTAACCTCCAGCTGCT 987
Db 1618 GGACAAAGTTTGAAGCTACGACCCCTTGCTCAGCAATTCATTAACCTCCAGCTGCT 1677
Oy 988 TCGATTTCAAAATTTTGGCAAAAGTGAAGATATCTTATTTGGGAATGGATGATGCTCAGAT 1047
Db 1678 TCGATTTCAAAATTTTGGCAAAAGTGAAGATATCTTATTTGGGAATGGATGATGCTCAGAT 1737
Oy 1048 GCACCTTGAAGACAGTTTCTATTCATATCAATTAATACATTTATGTTCCCAAGTTGTTAT 1107
Db 1738 GCACCTTGAAGACAGTTTCTATTCATATCAATTAATACATTTATGTTCCCAAGTTGTTAT 1797
Oy 1108 GCGGAAAAGTATATTTGGAACTTGTCCGCCAGATGAGTTCAAGAAACATTCAGATGG 1167
Db 1798 GCGGAAAAGTATATTTGGAACTTGTCCGCCAGATGAGTTCAAGAAACATTCAGATGG 1857
Oy 1168 GATGAATACTTGGAGAAAGATCTGCGAAGACCTTACCGCTTGACTTGTTCAGACCAATG 1227
Db 1858 GATGAATACTTGGAGAAAGATCTGCGAAGACCTTACCGCTTGACTTGTTCAGACCAATG 1917
Oy 1228 CCTTCCCAAGAGAGATTAGCAAAATTTAAGTAATTTCTGATTTCCAAACGGGT 1280
Db 1918 CCTTCCCAAGAGAGATTAGCAAAATTTAAGTAATTTCTGATTTCCAAACGGGT 1970

RESULT 8
US-09-864-761-31987/C
; Sequence 31987, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1
```

```

; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 31987
; LENGTH: 280
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009405.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
; OTHER INFORMATION: EST HUMAN HIT: BE535426.1, EVALUE 2.00e-32
; OTHER INFORMATION: NT HIT: D84430.1, EVALUE 1.00e-32
; OTHER INFORMATION: SWISSPROT HIT: P15624, EVALUE 6.00e-08
US-09-864-761-31987

Query Match 1.3%; Score 20; DB 9; Length 280;
Best Local Similarity 100.0%; Pred. No. 8.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1387 GATGAGCAATTGTGACT 1406
Db 156 GATGAGCAATTGTGACT 137
```

```

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15473
; LENGTH: 425
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009405.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
US-09-864-761-15473

Query Match 1.3%; Score 20; DB 9; Length 425;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1387 GATGAGCAATTGTGACT 1406
Db 391 GATGAGCAATTGTGACT 372
```

RESULT 10

```

US-10-136-224-2
; Sequence 2, Application US/10136224
; Publication No. US20020174448A1
; GENERAL INFORMATION:
; APPLICANT: AUERNHAMMER, CHRISTOPH J.
; APPLICANT: MELMED, SHLOMO
; TITLE OF INVENTION: SUPPRESSOR OF CYTOKINE SIGNALING
; TITLE OF INVENTION: (SOCs)-3 PROMOTER AND METHODS FOR ITS USE IN GENETIC THERAPY
```

```

: TITLE OF INVENTION: IN HUMANS
: FILE REFERENCE: P07 42591
: CURRENT APPLICATION NUMBER: US/10/136,224
: CURRENT FILING DATE: 2002-04-29
: PRIOR APPLICATION NUMBER: US/09/327,138
: PRIOR FILING DATE: 1999-06-07
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 2
: LENGTH: 2863
: TYPE: DNA
: ORGANISM: MURINE
: FEATURE:
: NAME/KEY: promoter
: LOCATION: (-2759)...(104)
: US-10-136-224-2

```

Query Match	1.3%	Score 20;	DB 13;	Length 2863;
Best Local Similarity	100.0%	Pred. No. 11;		
Matches 20;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	504	AGGSAAGGCTCTCCTAAGCA	523
Db	18	AGGSAAGGCTCTCCTAAGCA	37

```

RESULT 11
US-10-136-224-3
; Sequence 3, Application US/10136224
; Publication No. US2002017448A1
; GENERAL INFORMATION:
; APPLICANT: AUERNHAMMER, CHRISTOPH J.
; APPLICANT: MELMED, SHLOMO
; TITLE OF INVENTION: SUPPRESSOR OF CYTOKINE SIGNALING
; TITLE OF INVENTION: (SOC)-3 PROMOTER AND METHODS FOR ITS USE IN GENETIC THERAPY
; TITLE OF INVENTION: IN HUMANS
; FILE REFERENCE: P07 42591
; CURRENT APPLICATION NUMBER: US/10/136,224
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US/09/327, 138
; PRIOR FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 3686
; TYPE: DNA
; ORGANISM: MURINE
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (-2759)...(927)
US-10-136-224-3

```

```
Query Match      1.3%; Score 20; DB 13; Length 3686;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

Qy	504	AGGGAAGGCTCTCCTAAGCA	523
Db	18	AGGGAAGGCTCTCCTAAGCA	37

```

: RESULT 12
: US-10-136-224-1
: Sequence 1, Application US/10136224
: Publication No. US20020174448A1
: GENERAL INFORMATION:
: APPLICANT: AUERNHAMMER, CHRISTOPH J.
: APPLICANT: MELMED, SHLOMO
: TITLE OF INVENTION: SUPPRESSOR OF CYTOKINE SIGNALING
: TITLE OF INVENTION: (SOC5)-3 PROMOTER AND METHODS FOR ITS USE IN GENETIC THERAPY
: TITLE OF INVENTION: IN HUMANS
: FILE REFERENCE: P07 42591
: CURRENT APPLICATION NUMBER: US/10/136, 224

```

```

1 CURRENT FILING DATE: 2002-04-29
2 PRIOR APPLICATION NUMBER: US/09/327,138
3 PRIOR FILING DATE: 1999-06-07
4 NUMBER OF SEQ ID NOS: 30
5 SOFTWARE: FastSeq for Windows Version 3.0
6 SEQ ID NO 1
7 LENGTH: 3940
8 TYPE: DNA
9 ORGANISM: MURINE
10 FEATURE:
11 NAME/KEY: promoter
12 LOCATION: (-2907)..(1033)
13 FEATURE:
14 NAME/KEY: promoter
15 LOCATION: (0)..(0)
16 FEATURE:
17 NAME/KEY: promoter
18 LOCATION: (0)..(0)
19 FEATURE:
20 NAME/KEY: promoter
21 LOCATION: (0)..(0)
22 FEATURE:
23 NAME/KEY: promoter
24 LOCATION: (0)..(0)
25 FEATURE:
26 NAME/KEY: promoter
27 LOCATION: (0)..(0)
28 FEATURE:
29 NAME/KEY: promoter
30 LOCATION: (0)..(0)
31 NAME/KEY: promoter
32 LOCATION: (0)..(0)
33 US-10-136-224-1

```

Query Match	1.3%	Score 20;	DB 13;	Length 3940;
Best Local Similarity	100.0%	Pred. No. 11;		
Matches	20;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

Qy	504	AGGGAAGGCTCTCTTAAGCA	523
Db	166	AGGGAAGGCTCTCTTAAGCA	185

```

RESULT 13
US-10-240-485-46
Sequence 46, Application US/10240485
Publication No. US20030148327A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OR INVENTION: Diagnosis of Diseases Associated with
TITLE OF INVENTION: Metastasis
FILE REFERENCE: 5013.1007
CURRENT APPLICATION NUMBER: US/10/240.485
CURRENT FILING DATE: 2002-10-02
PRIOR APPLICATION NUMBER: PCT/EP01/03970
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 202
SEQ ID NO 46
LENGTH: 11996
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens
US-10-240-485-46

```

Query Match 1.3%; Score 20; DB 12; Length 11996;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 428 ATATGTAATATATGTAAT 447
|||||
Db 10139 ATATGTAATATATGTAAT 10158

RESULT 14
US-10-240-453-6/c
; Sequence 6, Application US/10240453
; Publication No. US20030148326A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPERBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
; TITLE OF INVENTION: Transcription
; TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associated
; FILE REFERENCE: 5013.1009
; CURRENT APPLICATION NUMBER: US/10/240,453
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03973
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 350
; SEQ ID NO 6
; LENGTH: 7450
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1729, 2292, 2294, 2295, 2299, 2798)
US-10-240-453-6

Query Match 1.3%; Score 19; DB 12; Length 7450;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1073 ATATCAATATATACATTTAT 1091
|||||
Db 6762 ATATCAATATATACATTTAT 6744

RESULT 15
US-10-195-144-87
; Sequence 87, Application US/10195144
; Publication No. US20030126646A1
; GENERAL INFORMATION:
; APPLICANT: BROWN, GREGORY G.
; APPLICANT: FORMANOVA, NATASA
; APPLICANT: DENDY, CHARLES
; APPLICANT: LANDRY, BENJOIT S.
; APPLICANT: CHEUNG, WING
; APPLICANT: JIN, HUA
; TITLE OF INVENTION: NUCLEAR FERTILITY RESTORER GENES AND METHODS OF USE IN
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 16313-0136
; CURRENT APPLICATION NUMBER: US/10/195,144
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/305,026
; PRIOR FILING DATE: 2001-07-12

; PRIOR APPLICATION NUMBER: 60/305,363
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/308,736
; PRIOR FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 271990
; TYPE: DNA
; ORGANISM: Raphanus sativum
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (144241)..(144300)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-195-144-87

Query Match 1.3%; Score 19; DB 14; Length 271990;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1395 ATTTGATGAAGTGTATGAT 1413
|||||
Db 34401 ATTTGATGAAGTGTATGAT 34419

Search completed: September 2, 2003, 20:40:03
Job time : 409 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2003, 15:09:29 ; Search time 454 Seconds
(without alignments)
8901.010 Million cell updates/sec

Title: US-09-872-523-6

Perfect score: 1497
Sequence: 1 atgtcgaattcttgaat.....cgaaaagracactatga 1497

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_19Jun03:*

```
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1497	100.0	1497	24	ABL52497	Caenorhabditis eleg
2	1446	96.6	1497	24	ABL52515	C. elegans mutant
3	1446	96.6	1497	24	ABL52516	C. elegans mutant
4	1446	96.6	1497	24	ABL52517	C. elegans mutant
5	833	55.6	2307	24	ABL52518	C. elegans lin-61
6	833	55.6	2307	24	ABL52519	C. elegans lin-61
7	833	55.6	2307	24	ABL52520	C. elegans lin-61
8	21	1.4	872	22	AAH04590	Human CDNA clone

9	21	1.4	2202	22	AAH17401	Human CDNA sequenc
10	20	1.3	280	22	ABA49041	Human breast cell
11	20	1.3	280	22	AAI21886	Probe #11819 for g
12	20	1.3	280	22	AAI07575	Probe #7566 used t
13	20	1.3	280	22	ABSI5121	Human genome-deriv
14	20	1.3	425	22	ABA43943	Human breast cell
15	20	1.3	425	22	AAI12703	Probe #2636 for ge
16	20	1.3	425	22	AAI02613	Probe #2604 used t
17	20	1.3	425	22	ABSO2626	Human genome-deriv
18	20	1.3	263	22	MAC85357	Human genome-deriv
19	20	1.3	3686	22	MAC83358	Murine SOCS-3 prom
20	20	1.3	3940	22	MAC83358	Murine SOCS-3 prom
21	20	1.3	11996	22	MAC85356	Murine SOCS-3 prom
22	19	1.3	2127	25	ABD51493	Human metastasis a
23	19	1.3	2178	25	ABX06430	Streptococcus pneu
24	19	1.3	2181	25	ABX42211	S. pneumoniae type
25	19	1.3	2883	21	AAI42886	Streptococcus pneu
26	19	1.3	3262	21	AAI42886	Streptococcus pneu
27	19	1.3	6870	23	ABLI17787	Arabidopsis thalia
28	19	1.3	7450	23	ABLI17786	Drosophila melanog
29	19	1.3	8494	19	AAV52296	Drosophila melanog
30	19	1.3	8753	24	ABL92270	DNA transcription
31	19	1.3	35048	23	ABL06102	Streptococcus pneu
32	19	1.3	2162598	25	ABX56454	Chemically treated
33	18	1.2	108	17	AAI12617	Drosophila melanog
34	18	1.2	286	24	ABX92085	Streptococcus pneu
35	18	1.2	302	24	ABX37370	U7.6 scfv PCR prim
36	18	1.2	360	24	ABX32053	Lung specific nucl
37	18	1.2	367	22	AAK76409	LR11-11 adenylate
38	18	1.2	409	22	AAK46491	Candida albicans e
39	18	1.2	512	24	ABQ57224	Human Immune/Haema
40	18	1.2	527	24	ABN76926	Human breast cell
41	18	1.2	539	23	ABV54427	Human colon cancer
42	18	1.2	742	21	AAZ98082	Human ORE1873 cDNA
43	18	1.2	742	22	AAI11695	Human prostate exp
44	18	1.2	742	22	AAI11695	Human secreted pro
45	18	1.2	781	21	AAZ97460	Human secreted pro

ALIGNMENTS

RESULT 1	ID	ABL52497 standard; DNA; 1497 BP.	ALIGNMENTS
XX	ABL52497		
XX	AC	ABL52497;	
XX	DT	17-JUL-2002 (first entry)	
XX	DE	Caenorhabditis elegans lin-61 nucleotide sequence SEQ ID NO:6.	
XX	DE	Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;	
XX	KW	cell proliferation; nematode; cancer; gene; ds.	
XX	OS	Caenorhabditis elegans.	
XX	FT	Key	Location/Qualifiers
XX	FT	CDS	1..1497
XX	FT		/tag= a
XX	FT		/product= "LIN-61 protein"
XX	PN	W0200194545-A2.	
XX	PD	13-DEC-2001.	
XX	PF	01-JUN-2001; 2001WO-US17909.	
XX	PR	02-JUN-2000; 2000US-208802P.	
XX	PA	(MAST) MASSACHUSETTS INST TECHNOLOGY.	
XX	PI	Horvitz HR, Davison EM, Lu X;	

Human CDNA sequenc
Human breast cell
Probe #11819 for g
Probe #7566 used t
Human genome-deriv
Human breast cell
Probe #2636 for ge
Probe #2604 used t
Human genome-deriv
Human genome-deriv
Murine SOCS-3 prom
Murine SOCS-3 prom
Murine SOCS-3 prom
Human metastasis a
Streptococcus pneu
S. pneumoniae type
Streptococcus pneu
Arabidopsis thalia
Drosophila melanog
Drosophila melanog
DNA transcription
Streptococcus pneu
Chemically treated
Drosophila melanog
Streptococcus pneu
U7.6 scfv PCR prim
Lung specific nucl
LR11-11 adenylate
Candida albicans e
Human Immune/Haema
Human breast cell
Human colon cancer
Human ORE1873 cDNA
Human prostate exp
Human secreted pro
Human secreted pro
Human prostate can

XX WPI: 2002-401590/43.
DR P-PSDB: ABB78650.
XX
XX New Caenorhabditis elegans nucleic acid involved in tumor suppressor
PT pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
PT diagnosing and treating cell proliferative diseases such as cancer
XX
XX Claim 24: Page 74: 116pp: English.
XX
XX The present invention describes a substantially pure nematode C. elegans
CC (Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide
CC (see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide
CC (see ABB78650), where the polypeptides comprise at least 130, 110, 130
CC contiguous amino acids of the 386, 322 and 498 amino acid sequences given
CC in ABB78648 to ABB78650, and modulate cell proliferation. (I) has
CC cytosolic activity, and can be used in gene therapy. (I) is useful for
CC modulating proliferation of a cell, and for identifying the compound that
CC modulates cell proliferation. (I) can be used for diagnosing an animal
CC (preferably, human) for the presence of the cell proliferation disease,
CC or an increased chance of developing the disease, by measuring LIN-8,
CC LIN-56 or LIN-61 nucleic acid expression in a sample obtained from the
CC animal, by measuring LIN-8, LIN-56 or LIN-61 polypeptide in the sample
CC using Ab or measuring amount of LIN-8, LIN-56, LIN-61 mRNA in the sample.
CC (I) is also useful for identifying the compound that modulates cell
CC proliferation. (I) and its mutants are useful for understanding cell
CC proliferative diseases including cancer, as well as in diagnosing and
CC treating cell proliferative diseases. (I) is also useful for identifying
CC tumour suppressors in other species such as mammals and may be used to
CC identify therapeutic compounds. The present sequence encodes C. elegans
CC LIN-61 from the present invention.
XX
XX Sequence 1497 BP; 497 A; 272 C; 312 G; 416 T; 0 other:
SQ
Query Match 100.0%; Score 1497; DB 24; Length 1497;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1497; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 541 GGACACGCTCTTGACATTAATTAATTCACATTTCTAGTAATAGCGGTAGCGCAAT 600
DB 541 GGACACGCTCTTGACATTAATTAATTCACATTTCTAGTAATAGCGGTAGCGCAAT 600
QY 601 CAAGAAATATGTGACGACGCAATGATATATCTATCAAGAAAGACCTTCCGATGCG 660
DB 601 CAAGAAATATGTGACGACGCAATGATATATCTATCAAGAAAGACCTTCCGATGCG 660
QY 661 CTTCAGATGACATGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 720
DB 661 CTTCAGATGACATGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 720
QY 721 GAGGAGACCTTCTCATATTTTCCGTGATGATTTGACAGATGATGATGATGATGAT 780
DB 721 GAGGAGACCTTCTCATATTTTCCGTGATGATTTGACAGATGATGATGATGATGAT 780
QY 781 GCGAAAAAGAAATATATTTGACGACGACGACGACGACGACGACGACGACGACGAC 840
DB 781 GCGAAAAAGAAATATATTTGACGACGACGACGACGACGACGACGACGACGACGAC 840
QY 841 AATCCAAATATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAT 900
DB 841 AATCCAAATATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAT 900
QY 901 CCCATGATTTGAGAAAAAGTTAAAGTTGGAACAAAGTTGAGCTCATGACCCCTGGCT 960
DB 901 CCCATGATTTGAGAAAAAGTTAAAGTTGGAACAAAGTTGAGCTCATGACCCCTGGCT 960
QY 961 CAGCAATTCATTAACCTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
DB 961 CAGCAATTCATTAACCTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 1021 CTATATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1021 CTATATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1081 AATCATTTATGTTCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
DB 1081 AATCATTTATGTTCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
QY 1141 GATGAGTTCAAGGAAACATTCAGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1141 GATGAGTTCAAGGAAACATTCAGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1201 CTACCGCTTGACTGTTCAACGCAATGCTTCCCAAGGAGATTAAGCAATTTAAGTA 1260
DB 1201 CTACCGCTTGACTGTTCAACGCAATGCTTCCCAAGGAGATTAAGCAATTTAAGTA 1260
QY 1261 AATCGATTTCCAAACGGGTGAGTACGCTGAGCTGCTGACATGCTGGAATATCAG 1320
DB 1261 AATCGATTTCCAAACGGGTGAGTACGCTGAGCTGCTGACATGCTGGAATATCAG 1320
QY 1321 TTTATTTGTCAGTACAGTGAATCACTTCATGGAAGACTGATTAATGCAATTTGAC 1380
DB 1321 TTTATTTGTCAGTACAGTGAATCACTTCATGGAAGACTGATTAATGCAATTTGAC 1380
QY 1381 GCGTGGATGAAGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
DB 1381 GCGTGGATGAAGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
QY 1441 GGATGCTGGAAGGCGACAGTATGCTTACAACTCCGAAAAAGTCAACTATGTA 1497
DB 1441 GGATGCTGGAAGGCGACAGTATGCTTACAACTCCGAAAAAGTCAACTATGTA 1497

RESULT 2
ABL52515
ID ABL52515 standard; DNA: 1497 BP.
XX
XX ABL52515;
AC
XX
XX 17-JUL-2002 (first entry)

XX DE C. elegans mutant lin-61 DNA sequence lin-61(n3446) SEQ ID NO:73.
 XX XX
 KM Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;
 XX cell proliferation; nematode; cancer; mutant; gene; ds.
 OS Caenorhabditis elegans.
 XX XX
 FH Key Location/Qualifiers
 FT CDS 1..1497
 FT /tag= a
 FT /product= "LIN-61(n3446) protein"
 PN WO200194545-A2.
 XX PD 13-DEC-2001.
 XX 01-JUN-2001; 2001WO-US17909.
 XX 02-JUN-2000; 2000US-208802P.
 XX (MASI) MASSACHUSETTS INST TECHNOLOGY.
 PA Horvitz HR, Davidson EM, Lu X;
 PI WPI; 2002-401590/43.
 DR P-PSDB; ABB78697.
 XX
 PT New Caenorhabditis elegans nucleic acid involved in tumor suppressor
 PT pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
 PT diagnosing and treating cell proliferative diseases such as cancer
 XX
 PS Claim 27: Page 112-113; 116pp; English.
 XX
 CC The present invention describes a substantially pure nematode C. elegans
 CC (Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide
 CC (see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide
 CC (see ABB78650), where the polypeptides comprise at least 130, 110, 130
 CC contiguous amino acids of the 386, 322 and 498 amino acid sequences given
 CC in ABB78648 to ABB78650, and modulate cell proliferation. (I) has
 CC cytostatic activity, and can be used in gene therapy. (I) is useful for
 CC modulating proliferation of a cell, and for identifying the compound that
 CC modulates cell proliferation. (I) can be used for diagnosing an animal
 CC (preferably, human) for the presence of the cell proliferation disease,
 CC or an increased chance of developing the disease, by measuring lin-8,
 CC lin-56 or lin-61 nucleic acid expression in a sample obtained from the
 CC animal, by measuring LIN-8, LIN-56 or LIN-61 polypeptide in the sample
 CC using Ab or measuring amount of lin-8, lin-56, lin-61 mRNA in the sample.
 CC (I) is also useful for identifying the compound that modulates cell
 CC proliferation. (I) and its mutants are useful for understanding cell
 CC proliferative diseases including cancer, as well as in diagnosing and
 CC treating cell proliferative diseases. (I) is also useful for identifying
 CC tumor suppressors in other species such as mammals and may be used to
 CC identify therapeutic compounds. The present sequence encodes a
 CC C. elegans mutant LIN-61 protein from the present invention.
 XX
 SQ Sequence 1497 BP; 497 A; 271 C; 312 G; 417 T; 0 other;

QY 181 ATCTTGAAGACAGTGGTCCATGATTATGACAAAGACCTGGATTCGATTCAGATGAG 240
 DB 181 ATCTTGAAGACAGTGGTCCATGATTATGACAAAGACCTGGATTCGATTCAGATGAG 240
 QY 241 TTTCGACGAAATGAAAGATTGGCGATACAGAGTTTGGCTCAGTTTATCGAGCTGAC 300
 DB 241 TTTCGACGAAATGAAAGATTGGCGATACAGAGTTTGGCTCAGTTTATCGAGCTGAC 300
 QY 301 ACGAAATTTGGCTCAATATTTTATTCGACGATATGTTTGGTTGGCAAGCCGCAATG 360
 DB 301 ACGAAATTTGGCTCAATATTTTATTCGACGATATGTTTGGTTGGCAAGCCGCAATG 360
 QY 361 AGTGATCCCAATATGATGATTAATGATGCTCCGCGCTTCGATCAACGAGAAATAC 420
 DB 361 AGTGATCCCAATATGATGATTAATGATGCTCCGCGCTTCGATCAACGAGAAATAC 420
 QY 421 CAAATGATATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
 DB 421 CAAATGATATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
 QY 481 TGGCTGTCTCCAAATTTGATGAGGAAAGGCTCTCTTAAGCAAGCATCGTTCAAGTT 540
 DB 481 TGGCTGTCTCCAAATTTGATGAGGAAAGGCTCTCTTAAGCAAGCATCGTTCAAGTT 540
 QY 541 GGACAACGCTTGAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
 DB 541 GGACAACGCTTGAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
 QY 601 CAGAAATATGTTGACGACGAGCAATGATATATACAAAGAAAGACTTCCCAATCG 660
 DB 601 CAGAAATATGTTGACGACGAGCAATGATATATACAAAGAAAGACTTCCCAATCG 660
 QY 661 CTTCCAGATGCAATGATGACGACGACGACGACGACGACGACGACGACGACGACGAC 720
 DB 661 CTTCCAGATGCAATGATGACGACGACGACGACGACGACGACGACGACGACGACGAC 720
 QY 721 GAGGAAAGCTTGTATATATTTCTGTTGATTTGACAGCAAGTAAATGATATCACTAAT 780
 DB 721 GAGGAAAGCTTGTATATATTTCTGTTGATTTGACAGCAAGTAAATGATATCACTAAT 780
 QY 781 GCGAAAGAGAAATATATGACGACACCAATTAATTTCTCAAGCAATTAATAATGAGCA 840
 DB 781 GCGAAAGAGAAATATATGACGACACCAATTAATTTCTCAAGCAATTAATAATGAGCA 840
 QY 841 AATCCAAAGTATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 900
 DB 841 AATCCAAAGTATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 900
 QY 901 CCCATGATTTGGAGAAAGATTAAAGTTGAGCAAAAGTTGAGCTCATGACCCCTTGCT 960
 DB 901 CCCATGATTTGGAGAAAGATTAAAGTTGAGCAAAAGTTGAGCTCATGACCCCTTGCT 960
 QY 961 CAGCAATTCATTAACCTCAGCTCGCTTGCATTCATTAATTTTGCAGAAATCTAAGAGAT 1020
 DB 961 CAGCAATTCATTAACCTCAGCTCGCTTGCATTCATTAATTTTGCAGAAATCTAAGAGAT 1020
 QY 1021 CTTATTTGGGAATGAGTGGTCCAGATGACGATGCAAGAGTTTCTATTTATATCAAT 1080
 DB 1021 CTTATTTGGGAATGAGTGGTCCAGATGACGATGCAAGAGTTTCTATTTATATCAAT 1080
 QY 1081 AATACATTTATGTTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 DB 1081 AATACATTTATGTTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 QY 1141 GATGAGTTCAAGGAACATTCAGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 DB 1141 GATGAGTTCAAGGAACATTCAGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 QY 1201 CTACCGCTTGAATTTGCAAGCAATGCTTCCCAAGAGATTTAGCAAAATTTAAGGTA 1260
 DB 1201 CTACCGCTTGAATTTGCAAGCAATGCTTCCCAAGAGATTTAGCAAAATTTAAGGTA 1260

QY 1261 ATCTGATTTCCAAACGGGTAGAGCTACGCCCTTGAGACTGCTGACATGTGTGAAATCAG 1320
DB 1261 ATCTGATTTCCAAACGGGTAGAGCTACGCCCTTGAGACTGCTGACATGTGTGAAATCAG 1320
QY 1321 TTTATTTTCACGCTACAGTGAATCATCTGATGGAACAGTGAATGATTAATTTGAC 1380
DB 1321 TTTATTTTCACGCTACAGTGAATCATCTGATGGAACAGTGAATGATTAATTTGAC 1380
QY 1381 GGCTGGATGGAAGAAATTTGATGACTGATGATGAGACTGCCATGATATTTCTACCGATA 1440
DB 1381 GGCTGGATGGAAGAAATTTGATGACTGATGATGAGACTGCCATGATATTTCTACCGATA 1440
QY 1441 GGATGCTGTGAAGCCGACAGTATGTCTACAACTCCGAAAAAGTACACATTTGA 1497
DB 1441 GGATGCTGTGAAGCCGACAGTATGTCTACAACTCCGAAAAAGTACACATTTGA 1497

RESULT 3
ABLS2516
ID ABL52516 standard; DNA; 1497 BP.
XX
AC ABL52516;
XX
DT 17-JUL-2002 (first entry)
XX
DE C. elegans mutant lin-61 DNA sequence lin-61(n3447) SEQ ID NO:74.
XX
XX Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;
KM cell proliferation; nematode; cancer; mutant; gene; ds.
XX
OS Caenorhabditis elegans.
XX
FH Key Location/Qualifiers
FT CDS 1..1497
FT /tag= a
FT /product= "LIN-61(n3447) protein"
XX
XX MO200194545-A2.
XX
PD 13-DEC-2001.
XX
PF 01-JUN-2001; 2001MO-US17909.
XX
PR 02-JUN-2000; 2000US-208802P.
XX
PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
PI Horvitz HR, Davison EM, Lu X;
XX
DR WPI: 2002-401590/43.
XX
DR P-PSDB; ABB78698.
XX
PT New Caenorhabditis elegans nucleic acid involved in tumor suppressor
PT pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
PT diagnosing and treating cell proliferative diseases such as cancer
XX
XX Claim 27; Page 113; 116pp; English.

CC The present invention describes a substantially pure nematode C. elegans
CC (Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide
CC (see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide
CC (see ABB78650), where the polypeptides comprise at least 130, 110, 130
CC contiguous amino acids of the 386, 322 and 498 amino acid sequences given
CC in ABB78648 to ABB78650, and modulate cell proliferation. (I) has
CC cytosolic activity, and can be used in gene therapy. (I) is useful for
CC modulating proliferation of a cell, and for identifying the compound that
CC modulates cell proliferation. (I) can be used for diagnosing an animal
CC (preferably, human) for the presence of the cell proliferation disease,
CC or an increased chance of developing the disease, by measuring lin-8,
CC lin-56 or lin-61 nucleic acid expression in a sample obtained from the
CC animal, by measuring LIN-8, LIN-56 or LIN-61 polypeptide in the sample
CC using Ab or measuring amount of lin-8, lin-56, lin-61 mRNA in the sample.
CC (I) is also useful for identifying the compound that modulates cell

CC proliferation. (I) and its mutants are useful for understanding cell
CC proliferative diseases including cancer, as well as in diagnosing and
CC treating cell proliferative diseases. (I) is also useful for identifying
CC tumour suppressors in other species such as mammals and may be used to
CC identify therapeutic compounds. The present sequence encodes a
CC C. elegans mutant LIN-61 protein from the present invention.

XX Sequence 1497 BP; 498 A; 272 C; 311 G; 416 T; 0 other;

QY Query Match 96.6%; Score 1446; DB 24; Length 1497;

DB Best Local Similarity 99.9%; Pred. No. 0;

Matches 1496; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTGAAATTTTCGAAAATTTGTCAGAGCTTAACAAAAATCGGACAGAAATCTGATTAAG 60

DB 1 ATGCTGAAATTTTCGAAAATTTGTCAGAGCTTAACAAAAATCGGACAGAAATCTGATTAAG 60

QY 61 ACCTACTTGGGGAAATCCTATTATCATCAGTGTGAGAAAGAAAAATCTTTCATTTCCA 120

DB 61 ACCTACTTGGGGAAATCCTATTATCATCAGTGTGAGAAAGAAAAATCTTTCATTTCCA 120

QY 121 GTTGAGACATTCATCAATCGTAACCTTACAGTTAATTTTAACGAATGCGTGAAGAGAGATT 180

DB 121 GTTGAGACATTCATCAATCGTAACCTTACAGTTAATTTTAACGAATGCGTGAAGAGAGATT 180

QY 181 ATCTTCGAACAGTGTGCTCATGATTATGACAAAGACTGCGATTGATTCACATCAGATGG 240

DB 181 ATCTTCGAACAGTGTGCTCATGATTATGACAAAGACTGCGATTGATTCACATCAGATGG 240

QY 241 TTTGACAGAAATTTGAAAAGTTGGGATACAGAGTTGCTGCTCAGTTATTCGAGCTGAC 300

DB 241 TTTGACAGAAATTTGAAAAGTTGGGATACAGAGTTGCTGCTCAGTTATTCGAGCTGAC 300

QY 301 ACGAAATTTTGGCTCAATATTTTATVCGAGAGATATGTTGGTTGGCAAAACCCGCAATG 360

DB 301 ACGAAATTTTGGCTCAATATTTTATVCGAGAGATATGTTGGTTGGCAAAACCCGCAATG 360

QY 361 AGTATCCCAATATGATGATAAATGTAATGCTCCGCGCTGCAATCAAGAAAGATAC 420

DB 361 AGTATCCCAATATGATGATAAATGTAATGCTCCGCGCTGCAATCAAGAAAGATAC 420

QY 421 CAAATGATATGTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 480

DB 421 CAAATGATATGTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 480

QY 481 TCGCTGTCTCCAAATTCGATGAAGGAGAGGCTCTCTAAGCAAGACTGCTTCAAAAGTT 540

DB 481 TCGCTGTCTCCAAATTCGATGAAGGAGAGGCTCTCTAAGCAAGACTGCTTCAAAAGTT 540

QY 541 GGACACGCTTGTGAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600

DB 541 GGACACGCTTGTGAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600

QY 601 CAAGAAATATGTGACGACGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 660

DB 601 CAAGAAATATGTGACGACGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 660

QY 661 CTTCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720

DB 661 CTTCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720

QY 721 GAGGGAACCTTCTCATATTTCTGTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 780

DB 721 GAGGGAACCTTCTCATATTTCTGTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 780

QY 781 GCGAAAAAGAAATATTTGATGACACACAAATTAATTTCTCAAGCAATTAATTAATTAAT 840

DB 781 GCGAAAAAGAAATATTTGATGACACACAAATTAATTTCTCAAGCAATTAATTAATTAATTA 840

QY 841 AATCCAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900

DB 841 AATCCAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900

```

QY 901 CCCATGATTTGGAGAAAGTTAAGTTGGACAAAGTTTGAGCTCATCGACCCCTTGCGT 960
DB 901 CCCATGATTTGGAGAAAGTTAAGTTGGACAAAGTTTGAGCTCATCGACCCCTTGCGT 960
QY 961 CAGCATTTCAATTAACCTCCACGCGCTTGATTTCTCAAAATTTTGGCAAACTGAGATAT 1020
DB 961 CAGCATTTCAATTAACCTCCACGCGCTTGATTTCTCAAAATTTTGGCAAACTGAGATAT 1020
QY 1021 CTATATGTGGAGATGATGCTCCAGATGACCTTGAAGACAGTTTCTTATTCATATCAAT 1080
DB 1021 CTATATGTGGAGATGATGCTCCAGATGACCTTGAAGACAGTTTCTTATTCATATCAAT 1080
QY 1081 AATPACATTTATGTTCCCACTGGTTATGCGGAAAGATATATTGGAACCTGTTCCGCCA 1140
DB 1081 AATPACATTTATGTTCCCACTGGTTATGCGGAAAGATATATTGGAACCTGTTCCGCCA 1140
QY 1141 GATGATTTCAAGGAACATTCAGATGGATGAAATTAATTTGGAGAAAGAAATCTGAGAAACC 1200
DB 1141 GATGATTTCAAGGAACATTCAGATGGATGAAATTAATTTGGAGAAAGAAATCTGAGAAACC 1200
QY 1201 CTACCGCTTGACTTGTTCAGCCCAATGCTTCCCAAGAGAGATTAGCAAAATTTAGATA 1260
DB 1201 CTACCGCTTGACTTGTTCAGCCCAATGCTTCCCAAGAGAGATTAGCAAAATTTAGATA 1260
QY 1261 ATTCTGATTTCCAAACGGGTAGGACTACCCCTTGAAGCTGCTGACATGTGTGAAAAATCAG 1320
DB 1261 ATTCTGATTTCCAAACGGGTAGGACTACCCCTTGAAGCTGCTGACATGTGTGAAAAATCAG 1320
QY 1321 TTTATTTGTCAGCTCACTGAATATGATTCATGGAAGACTGATTAATGTCATTTTCGAC 1380
DB 1321 TTTATTTGTCAGCTCACTGAATATGATTCATGGAAGACTGATTAATGTCATTTTCGAC 1380
QY 1381 GGCTGGGATGAAGAAATTTGATGAACTGTATGATGTGATGCCATGATATTTACCGGATA 1440
DB 1381 GGCTGGGATGAAGAAATTTGATGAACTGTATGATGTGATGCCATGATATTTACCGGATA 1440
QY 1441 GGATGGTGTGAAGCGCACAGTTATGTTTCAACAACCTCCGAAAAAGTACAACTATTTGA 1497
DB 1441 GGATGGTGTGAAGCGCACAGTTATGTTTCAACAACCTCCGAAAAAGTACAACTATTTGA 1497

RESULT 4
ABL52517
ID ABL52517 standard; DNA: 1497 BP.
XX
AC ABL52517;
XX
DT 17-JUL-2002 (first entry)
XX
DE C. elegans mutant lin-61 DNA sequence lin-61(n3624) SEQ ID NO:75.
XX
KW Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;
XX cell proliferation; nematode; cancer; mutant; gene; ds.
XX
OS Caenorhabditis elegans.
XX
FH key Location/Qualifiers
FT CDS 1..1497
FT /tag= a
FT /product= "LIN-61(n3624) protein"
XX
PN WO200194545-A2.
XX
PD 13-DEC-2001.
XX
PF 01-JUN-2001: 2001MO-US17909.
XX
PR 02-JUN-2000: 2000US-208802P.
XX
PA (MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX
PI Horvitz HR, Davison EM, Lu X:

```

```

DR WPI: 2002-401590/43.
DR P-SDS: ABB78699.
XX
PT New Caenorhabditis elegans nucleic acid involved in tumor suppressor
PT pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
PT diagnosing and treating cell proliferative diseases such as cancer
XX
PS Claim 27: Page 113-114: 116pp; English.
XX
CC The present invention describes a substantially pure nematode C. elegans
CC (Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide
CC (see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide
CC (see ABB78650), where the polypeptides comprise at least 130, 110, 130
CC contiguous amino acids of the 386, 322 and 498 amino acid sequences given
CC in ABB78648 to ABB78650, and modulate cell proliferation. (I) has
CC cytostatic activity, and can be used in gene therapy. (I) is useful for
CC modulating proliferation of a cell, and for identifying the compound that
CC modulates cell proliferation. (I) can be used for diagnosing an animal
CC (preferably, human) for the presence of the cell proliferation disease,
CC or an increased chance of developing the disease, by measuring lin-8,
CC lin-56 or lin-61 nucleic acid expression in a sample obtained from the
CC animal, by measuring LIN-8, LIN-56 or LIN-61 polypeptide in the sample.
CC using Ab or measuring amount of lin-8, lin-56, lin-61 mRNA in the sample.
CC (I) is also useful for identifying the compound that modulates cell
CC proliferation. (I) and its mutants are useful for understanding cell
CC proliferative diseases including cancer, as well as in diagnosing and
CC treating cell proliferative diseases. (I) is also useful for identifying
CC tumour suppressors in other species such as mammals and may be used to
CC identify therapeutic compounds. The present sequence encodes a
CC C. elegans mutant LIN-61 protein from the present invention.
XX
SQ Sequence 1497 BP; 497 A; 271 C; 312 G; 417 T; 0 other:
Query Match 96.6%; Score 1446; DB 24; Length 1497;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1496; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGCTGAAATTTCTGAAAATTTGTCAGAGCTAACCAAAAATCGGACAGAAATCTGATAAG 60
DB 1 ATGCTGAAATTTCTGAAAATTTGTCAGAGCTAACCAAAAATCGGACAGAAATCTGATAAG 60
QY 61 ACCTACTTTGGGAATTCCTATTACATCACTTGTGAGAAAGAAATCTTTCATTTCCA 120
DB 61 ACCTACTTTGGGAATTCCTATTACATCACTTGTGAGAAAGAAATCTTTCATTTCCA 120
QY 121 GTTGAAGCAATTCATGTAACCTTACAGTTAATTTTACGAATGCGTGAAGAGAGTT 180
DB 121 GTTGAAGCAATTCATGTAACCTTACAGTTAATTTTACGAATGCGTGAAGAGAGTT 180
QY 181 ATCTTGAAAACAGTGGTCCATGATTTATGACAAAGACGCGATTCGATTCAGATGAG 240
DB 181 ATCTTGAAAACAGTGGTCCATGATTTATGACAAAGACGCGATTCGATTCAGATGAG 240
QY 241 TTTGACAGAAATTAAGAAAGTTTGGCATACAGAGTTCTGGCTCAGTTATCGGAGCTGAC 300
DB 241 TTTGACAGAAATTAAGAAAGTTTGGCATACAGAGTTCTGGCTCAGTTATCGGAGCTGAC 300
QY 301 ACGAAATTTTGGCTCAATATTTTATCGGACGATATGTTGGTTGGCAAAAGCGCCGAATG 360
DB 301 ACGAAATTTTGGCTCAATATTTTATCGGACGATATGTTGGTTGGCAAAAGCGCCGAATG 360
QY 361 AGTGATCCCAATATGATTAATTTGATATGCTCCGCCGTTGCAATCAACGAAGATAC 420
DB 361 AGTGATCCCAATATGATTAATTTGATATGCTCCGCCGTTGCAATCAACGAAGATAC 420
QY 421 CAAAATGATATGTAATTAATTAATTAATTTGATGATGCGCAAAATCGTGGCCAAACT 480
DB 421 CAAAATGATATGTAATTAATTAATTAATTTGATGATGCGCAAAATCGTGGCCAAACT 480
QY 481 TCGCTGTCCAAAATTCGATGAAGGAGGCTCTCTTAAGCAAGCATGTTCAAAGTT 540
DB 481 TCGCTGTCCAAAATTCGATGAAGGAGGCTCTCTTAAGCAAGCATGTTCAAAGTT 540

```

QY	541	GGCAACAGCTGTGACCAATTAATTAATTCATCTCAATTCAGAAATPACCGATGAGCGCAAT	600
Db	541	GGCAACAGCTGTGACCAATTAATTAATTCATCTCAATTCAGAAATPACCGATGAGCGCAAT	600
QY	601	CAAGAAATATGTGAGACGACGAATGAATGAATATCTATCAACAAGAAAGACCTTCCGGAATCG	660
Db	601	CAAGAAATATGTGAGACGACGAATGAATGAATATCTATCAACAAGAAAGACCTTCCGGAATCG	660
QY	661	CTTCCAGATGCAGATGACGACGACGACAGCAAGCTTTAGCTCTGGATCTCAATATTTGGATAGAC	720
Db	661	CTTCCAGATGCAGATGACGACGACGACGACAGCAAGCTTTAGCTCTGGATCTCAATATTTGGATAGAC	720
QY	721	GAGGAAAGCTCTTCATATTTCCGTGTGGAAATTTGCACACGCAATGGATATCAACTAAAT	780
Db	721	GAGGAAAGCTCTTCATATTTCCGTGTGGAAATTTGCACACGCAATGGATATCAACTAAAT	780
QY	781	GGCAAAAAAGCAATATATTGAGACACACCAATTAATTTCTCAAGCAATTAATAAATGAGAA	840
Db	781	GGCAAAAAAGCAATATATTGAGACACACCAATTAATTTCTCAAGCAATTAATAAATGAGAA	840
QY	841	AATCCAGATATGACTCAGACGACGCTCAGCATTTGATCAATTTAGCAAAAAGATCCAAATGAT	900
Db	841	AATCCAGATATGACTCAGACGACGCTCAGCATTTGATCAATTTAGCAAAAAGATCCAAATGAT	900
QY	901	CCCATGATTTGGAAAAAGTTAAGTTGAGCAAAAAGTTTGAGCTCATCGACCCCTTGCT	960
Db	901	CCCATGATTTGGAAAAAGTTAAGTTGAGCAAAAAGTTTGAGCTCATCGACCCCTTGCT	960
QY	961	CAGCAATTCATAATACCTCCACGCTGGCTTCGATTCCAATTTTTCGCAAAACCTGAAAGATAT	1020
Db	961	CAGCAATTCATAATACCTCCACGCTGGCTTCGATTCCAATTTTTCGCAAAACCTGAAAGATAT	1020
QY	1021	CTTATTTGGGAATGATGGTCCGAGATGCACTTGAAGACGTTTCCATTTCAATATCAAT	1080
Db	1021	CTTATTTGGGAATGATGGTCCGAGATGCACTTGAAGACGTTTCCATTTCAATATCAAT	1080
QY	1081	AATACATTTATGTTCCCAAGTTAGTGGTATGCGGAAAAAGTATATTTGGAACCTGTTCCGCA	1140
Db	1081	AATACATTTATGTTCCCAAGTTAGTGGTATGCGGAAAAAGTATATTTGGAACCTGTTCCGCA	1140
QY	1141	GATGAGTTCAAAGGACATTCAGATGGGATGAATACCTTGGAAAAAGATCTGCAGAAACC	1200
Db	1141	GATGAGTTCAAAGGACATTCAGATGGGATGAATACCTTGGAAAAAGATCTGCAGAAACC	1200
QY	1201	CTACCGCTTGACTTGTTCBAAGCCAATGCTTCCCAAGAGATTTAGCAAAATTTTAAGSTA	1260
Db	1201	CTACCGCTTGACTTGTTCBAAGCCAATGCTTCCCAAGAGATTTAGCAAAATTTTAAGSTA	1260
QY	1261	ATTCGATTTCCAAACGGGTAGGACATGACGCTTGAAGCTGCTGCATGTGTGAAAAATCAG	1320
Db	1261	ATTCGATTTCCAAACGGGTAGGACATGACGCTTGAAGCTGCTGCATGTGTGAAAAATCAG	1320
QY	1321	TTTATTTTGTCCAGCTACAGTAAATCAGTTTCATGGAAGACTGATTAATGTCAATTTTGCAC	1380
Db	1321	TTTATTTTGTCCAGCTACAGTAAATCAGTTTCATGGAAGACTGATTAATGTCAATTTTGCAC	1380
QY	1381	GGCTGGGATGAAGATTTGATGAACCTGTATGATGTGGACTCCCATGATATTTCTCCACGATA	1440
Db	1381	GGCTGGGATGAAGATTTGATGAACCTGTATGATGTGGACTCCCATGATATTTCTCCACGATA	1440
QY	1441	GGATGCGTGAAGGCGACAGTTATGTTCTACAACTCCGAAAAAAGTACAACTATTTGA	1497
Db	1441	GGATGCGTGAAGGCGACAGTTATGTTCTACAACTCCGAAAAAAGTACAACTATTTGA	1497

RESULT	5
ABL52518	
ID	ABL52518 standard; DNA; 2307 BP
XX	
AC	ABL52518:
XX	
DT	17-JUL-2002 (first entry)
XX	

Query Match	Best Local Similarity	55.6%	Score 833	DB 24	Length 2307
Matches 833	Conservative 0	Mismatches 0	Indels 0	Gaps 0	
448	AATTCATGTTGGCGAAATGTCGGCCAAACTTCGCTCCAAATTCGATGAAGG	507			
1138	AATTCATGTTGGCGAAATGTCGGCCAAACTTCGCTCCAAATTCGATGAAGG	1197			
508	AAGGCTCTCTACGACAGCATCGTTTCAAGTTGGACAACTGTTGAACATTAATAT	567			
1198	AAGGCTCTCTACGACAGCATCGTTTCAAGTTGGACAACTGTTGAACATTAATAT	1257			
568	TCCATATCTACTGAATACGGCGTACGGCGAATTCAGAAATATGTGACGACGAATGAT	627			
1258	TCCATATCTACTGAATACGGCGTACGGCGAATTCAGAAATATGTGACGACGAATGAT	1317			
628	GTATCTATCACAAAGAAAGACCTTCGCCAATCGCTCCAGATGCGAGACAGACAA	687			
1318	GTATCTATCACAAAGAAAGACCTTCGCCAATCGCTCCAGATGCGAGACAGACAA	1377			
688	GTCCTTACCTGATGTCATATTTGGATACGAGGACGACCTTCTATATTTCTCTGT	747			

Db 1378 GTCCTTAAGCTCTGATCTCATATATGATAGACGAGGAACTTCTTCATATATTCCTGTT 1437
QY 748 GGATTTGACAGCATCATATGATATCACTAAATCGAAGAAAGATATATTTAGACACACA 807
Db 1438 GGATTTGACAGCATCATATGATATCACTAAATCGAAGAAAGATATATTTAGACACACA 1497
QY 808 AATTAATTTGCTCAAGCAATTAATAAATGAGAGAAATCCAGATATGACTCAGACAGCTC 867
Db 1498 AATTAATTTGCTCAAGCAATTAATAAATGAGAGAAATCCAGATATGACTCAGACAGCTC 1557
QY 868 ACATTTATCATATTAAGCAAAAGATCCATTTGATTTGAGAGAAAGTTAAGCTT 927
Db 1558 ACATTTATCATATTAAGCAAAAGATCCATTTGATTTGAGAGAAAGTTAAGCTT 1617
QY 928 GGACAAAGTTTGAGCTCATGAGCCCTTGCTGAGCAATTCATTAACCTCCAGCTGCT 987
Db 1618 GGACAAAGTTTGAGCTCATGAGCCCTTGCTGAGCAATTCATTAACCTCCAGCTGCT 1677
QY 988 TCGATTTCTCAAAATTTTGCAAAAGTGAAGATATCTTATTTGGGAATGATGTCAGAT 1047
Db 1678 TCGATTTCTCAAAATTTTGCAAAAGTGAAGATATCTTATTTGGGAATGATGTCAGAT 1737
QY 1048 GCATTTGAAGCACTTTTCCATTCATATCATATTAATACATTTATGTTCCAGTTGTTAT 1107
Db 1738 GCATTTGAAGCACTTTTCCATTCATATCATATTAATACATTTATGTTCCAGTTGTTAT 1797
QY 1108 GCGGAAAGATATATTTGGAACCTGTTCCGCGAGATAGTTCAAGAGCAATTCAGATNG 1167
Db 1798 GCGGAAAGATATATTTGGAACCTGTTCCGCGAGATAGTTCAAGAGCAATTCAGATNG 1857
QY 1168 GATGAATTAAGTGAAGAAAGATCTGCAAGAAACCTTACCGCTTGACTTGTTCAGCAATG 1227
Db 1858 GATGAATTAAGTGAAGAAAGATCTGCAAGAAACCTTACCGCTTGACTTGTTCAGCAATG 1917
QY 1228 CCTTCCCAAGAGATATGACAAATTTAAGGTAATTTGATTTCCAAACGGGT 1280
Db 1918 CCTTCCCAAGAGATATGACAAATTTAAGGTAATTTGATTTCCAAACGGGT 1970

RESULT 6
ABL52519
ID ABL52519 standard; DNA: 2307 BP.
AC ABL52519;
XX
XX 17-JUL-2002 (first entry)
XX
XX C. elegans lin-61(sy223) DNA sequence SEQ ID NO:77.
XX
XX Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;
XX cell proliferation; nematode; cancer; mutant; gene; ds.
XX
XX Caenorhabditis elegans.
XX
XX WO200194545-A2.
XX
XX 13-DEC-2001.
XX
XX 01-JUN-2001; 2001WO-US17909.
XX
XX 02-JUN-2000; 2000US-208802P.
XX
XX (MASI) MASSACHUSETTS INST TECHNOLOGY.
XX
XX Horvitz HR, Davison EM, Lu X;
XX
XX WPI: 2002-401590/43.
XX
XX New Caenorhabditis elegans nucleic acid involved in tumor suppressor
XX pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
XX diagnosing and treating cell proliferative diseases such as cancer
XX
XX Disclosure; Page 115; 116pp; English.

XX
CC The present invention describes a substantially pure nematode C. elegans
CC (Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide
CC (see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide
CC (see ABB78650), where the polypeptides comprise at least 130, 110, 130
CC continuous amino acids of the 386, 322 and 498 amino acid sequences given
CC in ABB78648 to ABB78650, and modulate cell proliferation. (I) has
CC cytosolic activity, and can be used in gene therapy. (I) is useful for
CC modulating proliferation of a cell, and for identifying the compound that
CC modulates cell proliferation. (I) can be used for diagnosing an animal
CC (preferably, human) for the presence of the cell proliferation disease,
CC or an increased chance of developing the disease, by measuring lin-8,
CC lin-56 or lin-61 nucleic acid expression in a sample obtained from the
CC animal, by measuring LIN-8, LIN-56 or LIN-61 polypeptide in the sample.
CC using Ab or measuring amount of lin-8, lin-56, lin-61 mRNA in the sample.
CC (I) is also useful for identifying the compound that modulates cell
CC proliferation. (I) and its mutants are useful for understanding cell
CC proliferative diseases including cancer, as well as in diagnosing and
CC treating cell proliferative diseases. (I) is also useful for identifying
CC tumour suppressors in other species such as mammals and may be used to
CC identify therapeutic compounds. The present sequence represents a
CC C. elegans lin-61(sy223) DNA sequence from the present invention.

XX
SQ Sequence 2307 BP; 741 A; 393 C; 465 G; 708 T; 0 other;

Query Match 55.6%; Score 833; DB 24; Length 2307;

Best Local Similarity 100.0%; Pred. No. 0; Matches: 833; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 448 AATTGATGATGAGCGGAATTCGTGGCCAAATCTGCTGCTCCAAATTTGATGAAGGG 507
Db 1138 AATTGATGATGAGCGGAATTCGTGGCCAAATCTGCTGCTCCAAATTTGATGAAGGG 1197
QY 508 AAGGCTCTCTAAGCAAGCATCGTTTCAAAAGTTGAGACAAGCTCTTGCACTTTAATAT 567
Db 1198 AAGGCTCTCTAAGCAAGCATCGTTTCAAAAGTTGAGACAAGCTCTTGCACTTTAATAT 1257
QY 568 TCCAAATCTACTGAATATACGCGTATGCGCAATTCAGAAATATGAGCAGCAATGAAT 627
Db 1258 TCCAAATCTACTGAATATACGCGTATGCGCAATTCAGAAATATGAGCAGCAATGAAT 1317
QY 1318 GTATCTATACAAAGAAAGACTTCCGGAATCGCTTCAGATGAGTGAAGCAGACAA 687
Db 1318 GTATCTATACAAAGAAAGACTTCCGGAATCGCTTCAGATGAGTGAAGCAGACAA 1377
QY 688 GTCCTTAGCTCTGATCTCATATATGATATGAGAGGAAAGCTTCTCATATTTCTGTT 747
Db 1378 GTCCTTAGCTCTGATCTCATATATGATATGAGAGGAAAGCTTCTCATATTTCTGTT 1437
QY 748 GGATTTGACAGCATCATATGATATCACTAAATCGAAGAAAGATATATTTAGACACACA 807
Db 1438 GGATTTGACAGCATCATATGATATCACTAAATCGAAGAAAGATATATTTAGACACACA 1497
QY 808 AATTAATTTGCTCAAGCAATTAATAAATGAGAGAAATCCAGATATGACTCAGACAGCTC 867
Db 1498 AATTAATTTGCTCAAGCAATTAATAAATGAGAGAAATCCAGATATGACTCAGACAGCTC 1557
QY 868 ACATTTATCATATTAAGCAAAAGATCCATTTGATTTGAGAGAAAGTTAAGCTT 927
Db 1558 ACATTTATCATATTAAGCAAAAGATCCATTTGATTTGAGAGAAAGTTAAGCTT 1617
QY 928 GGACAAAGTTTGAGCTCATGAGCCCTTGCTGAGCAATTCATTAACCTCCAGCTGCT 987
Db 1618 GGACAAAGTTTGAGCTCATGAGCCCTTGCTGAGCAATTCATTAACCTCCAGCTGCT 1677
QY 988 TCGATTTCTCAAAATTTTGCAAAAGTGAAGATATCTTATTTGGGAATGATGTCAGAT 1047
Db 1678 TCGATTTCTCAAAATTTTGCAAAAGTGAAGATATCTTATTTGGGAATGATGTCAGAT 1737
QY 1048 GCATTTGAAGCACTTTTCCATTCATATCATATTAATACATTTATGTTCCAGTTGTTAT 1107
Db 1738 GCATTTGAAGCACTTTTCCATTCATATCATATTAATACATTTATGTTCCAGTTGTTAT 1797


```

XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX 27-AUG-1999; 99JP-0300253.
XX 11-JAN-2000; 2000JP-0118776.
XX 02-MAY-2000; 2000JP-0183767.
XX 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX
XX Claim 1; SEQ ID 1425; 2537bp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in
XX the specification. The primer sets can be used in antisense therapy and
XX in gene therapy. The primers are useful for synthesizing polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
XX AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
XX represent oligonucleotides, all of which are used in the exemplification
XX of the present invention.
XX
XX Sequence 872 BP; 285 A; 138 C; 177 G; 263 T; 9 other;
XX
XX Query Match 1.4%; Score 21; DB 22; Length 872;
XX Best Local Similarity 100.0%; Pred. No. 4.9;
XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 873 TGAATCAATTAGCAAAAGATCC 893
XX |
XX 747 TGAATCAATTAGCAAAAGATCC 767
XX
XX RESULT 9
XX AAH17401
XX ID AAH17401 standard; cDNA; 2202 BP.
XX
XX AC AAH17401;
XX
XX 26-JUN-2001 (first entry)
XX
XX Human cDNA sequence SEQ ID NO:16839.
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX

```

```

OS Homo sapiens.
XX
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX 27-AUG-1999; 99JP-0300253.
XX 11-JAN-2000; 2000JP-0118776.
XX 02-MAY-2000; 2000JP-0183767.
XX 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX
XX Claim 8; SEQ ID 16839; 2537bp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in
XX the specification. The primer sets can be used in antisense therapy and
XX in gene therapy. The primers are useful for synthesizing polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
XX AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
XX represent oligonucleotides, all of which are used in the exemplification
XX of the present invention.
XX
XX Sequence 2202 BP; 703 A; 416 C; 488 G; 595 T; 0 other;
XX
XX Query Match 1.4%; Score 21; DB 22; Length 2202;
XX Best Local Similarity 100.0%; Pred. No. 5;
XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 873 TGAATCAATTAGCAAAAGATCC 893
XX |
XX 746 TGAATCAATTAGCAAAAGATCC 766
XX
XX RESULT 10
XX ABA49041/C
XX ID ABA49041 standard; DNA; 280 BP.
XX
XX AC ABA49041;
XX
XX 01-FEB-2002 (first entry)
XX
XX Human breast cell single exon nucleic acid probe #7736.
XX
XX Human; microarray; single exon probe; gene expression; breast;
XX

```


PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-476286/51.
XX
XX Novel single exon nucleic acid probe used to measuring gene expression
PT in a human breast -
XX
XX Claim 25; SEQ ID No 7566; 322pp; English.
XX
XX The present invention relates to novel single exon nucleic acid probes.
CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridizes at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and
CC non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 280 BP; 78 A; 53 C; 43 G; 106 T; 0 other;
SQ
Query Match 1.3%; Score 20; DB 22; Length 280;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1387 GATGAGAAATTGATGAACT 1406
Db 156 GATGAGAAATTGATGAACT 137
|||||
RESULT 13
ABS15121/c
ID ABS15121 standard; DNA: 280 BP.
XX
XX ABS15121:
AC
XX
XX 19-AUG-2002 (first entry)
DE Human genome-derived single exon probe ORF from lung SEQ ID No 15112.
XX
XX Human; ds: single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease; open reading frame; ORF.
XX
XX Homo sapiens.
OS
XX
XX MO200186003-A2.
PN
XX
XX 15-NOV-2001.
PD
XX
XX 30-JAN-2001; 2001WO-0500665.
PE
XX
XX 04-FEB-2000; 2000US-180312P.
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR

PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2002-114183/15.
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples -
XX
XX Claim 4; SEQ ID No 15112; 634pp; English.
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of
CC probes: the novel set of probes which hybridize at high stringency to a
CC nucleic acid expressed in the human lung; measuring gene expression in a
CC sample derived from human lung, comprising (a) contacting the array with
CC a collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a single exon
CC probe open reading frame of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 280 BP; 78 A; 53 C; 43 G; 106 T; 0 other;
SQ
Query Match 1.3%; Score 20; DB 24; Length 280;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1387 GATGAGAAATTGATGAACT 1406
Db 156 GATGAGAAATTGATGAACT 137
|||||
RESULT 14
ABA43943/c
ID ABA43943 standard; DNA: 425 BP.
XX
XX ABA43943:
AC
XX

```

DT 01-FEB-2002 (first entry)
XX
DE Human breast cell single exon nucleic acid probe #2638.
XX
KW Human: microarray; single exon probe; gene expression; breast;
XX disease; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PE 30-JAN-2001; 2001WO-US00662.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes,
XX useful for measuring gene expression in sample derived from human
XX breast, comprises number of single exon nucleic acid probes -
XX
PS Claim 1; SEQ ID NO 2638; 327bp + sequence listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human breast and BT 474 cells. The method involves contacting
XX the probes with a collection of detectably labelled nucleic acids
XX derived from mRNA of human breast, and then measuring the label
XX bound to each probe of the microarray. The probes are useful for
XX verifying the expression of regions of genomic DNA predicted to
XX encode proteins. They are useful for gene discovery, and for
XX determining predisposition and/or prognosing breast disease. Gene
XX expression analysis is useful for assessing the toxicity of chemical
XX agents on cells. The microarray of this invention presents a far greater
XX diversity of probes for measuring gene expression, with far less bias
XX than expressed sequence tag microarrays. The method is suitable for
XX rapid production of functional information from genomic sequence. The
XX present sequence is a single exon nucleic acid probe of the invention.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 425 BP; 122 A; 88 C; 66 G; 149 T; 0 other;
XX
Query Match 1.3%; Score 20; DB 22; Length 425;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1387 GATGAGAAATTGATGAACT 1406
DB 391 GATGAGAAATTGATGAACT 372
XX
RESULT 15
AA112703/c
ID AA112703 standard; DNA: 425 BP.
XX
AC AA112703;
XX
DT 12-OCT-2001 (first entry)
XX

```

```

DE Probe #2636 for gene expression analysis in human cervical cell sample.
XX
XX Probe: human; microarray; gene expression; cervical epithelial cell;
XX cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PE 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human cervical epithelial cells -
XX
PS Claim 25; SEQ ID NO 2636; 487bp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
XX (SENP). The present sequence is one such probe. The SENPs are derived
XX from human HeLa cells. The SENPs can be used to produce a single exon
XX microarray, which can be used for measuring human gene expression in a
XX sample derived from human cervical epithelial cells. By measuring gene
XX expression, the probes are therefore useful in grading and/or staging
XX of diseases of the cervix, notably cervical cancer.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 425 BP; 122 A; 88 C; 66 G; 149 T; 0 other;
XX
Query Match 1.3%; Score 20; DB 22; Length 425;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1387 GATGAGAAATTGATGAACT 1406
DB 391 GATGAGAAATTGATGAACT 372
XX
Search completed: September 2, 2003, 17:58:31
Job time : 460 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2003, 17:43:50 ; Search time 107 Seconds
(Without alignments)
6175.239 Million cell updates/sec

Title: US-09-872-523-6

Perfect score: 1497
Sequence: 1 atgtcgaattctgaat.....cgaagaaagacacactatga 1497

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 569978 seqs, 220691566 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued_Patents_NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	1.3	2044	4 US-09-327-138C-38	Sequence 38, Appl
2	20	1.3	2863	4 US-09-327-138C-2	Sequence 2, Appl1
3	20	1.3	3646	4 US-09-327-138C-36	Sequence 36, Appl
4	20	1.3	3665	4 US-09-327-138C-35	Sequence 35, Appl
5	20	1.3	3686	4 US-09-327-138C-3	Sequence 3, Appl1
6	20	1.3	3686	4 US-09-327-138C-37	Sequence 37, Appl
7	20	1.3	3940	4 US-09-327-138C-1	Sequence 1, Appl1
8	19	1.3	8494	4 US-08-961-527-163	Sequence 163, Appl
9	18	1.2	742	4 US-09-482-273-76	Sequence 76, Appl
10	18	1.2	781	4 US-09-328-475C-321	Sequence 321, Appl
11	18	1.2	1172	4 US-08-936-165A-8	Sequence 8, Appl1
12	17	1.1	921	4 US-08-795-475-2	Sequence 2, Appl1
13	17	1.1	1005	3 US-09-446-504-79	Sequence 79, Appl
14	17	1.1	1005	4 US-09-712-266-79	Sequence 79, Appl
15	17	1.1	1050	4 US-09-107-532A-2184	Sequence 2184, Appl
16	17	1.1	1189	3 US-09-497-779A-7	Sequence 7, Appl1
17	17	1.1	1251	3 US-09-355-115-1	Sequence 1, Appl1
18	17	1.1	1305	4 US-09-329-234A-6	Sequence 6, Appl1
19	17	1.1	1308	2 US-08-795-475-4	Sequence 4, Appl1
20	17	1.1	1316	4 US-09-370-838-30	Sequence 30, Appl
21	17	1.1	1618	1 US-07-726-607C-2	Sequence 2, Appl1
22	17	1.1	1618	1 US-07-843-949A-2	Sequence 2, Appl1
23	17	1.1	1618	1 US-08-218-978-2	Sequence 2, Appl1
24	17	1.1	1782	4 US-09-220-132-158	Sequence 158, Appl
25	17	1.1	2207	4 US-09-254-504-6	Sequence 6, Appl1
26	17	1.1	2228	1 US-07-726-607C-1	Sequence 1, Appl1
27	17	1.1	2228	1 US-07-843-949A-1	Sequence 1, Appl1

28	17	1.1	2228	2 US-08-218-978-1	Sequence 1, Appl1
29	17	1.1	2389	3 US-08-961-083-55	Sequence 55, Appl
30	17	1.1	2389	4 US-09-536-784-55	Sequence 55, Appl
31	17	1.1	2451	4 US-09-468-656A-9	Sequence 9, Appl1
32	17	1.1	2634	3 US-08-949-386-26	Sequence 26, Appl
33	17	1.1	2634	3 US-08-450-562-26	Sequence 26, Appl
34	17	1.1	2634	4 US-08-984-709A-26	Sequence 26, Appl
35	17	1.1	2634	4 US-08-450-272-26	Sequence 26, Appl
36	17	1.1	2708	1 US-08-384-556A-10	Sequence 10, Appl
37	17	1.1	2708	1 US-08-331-355A-25	Sequence 25, Appl
38	17	1.1	2708	5 PCT-US94-12364-25	Sequence 25, Appl
39	17	1.1	2708	5 PCT-US95-07753-10	Sequence 10, Appl
40	17	1.1	2712	3 US-08-949-386-38	Sequence 38, Appl
41	17	1.1	2712	3 US-08-450-562-38	Sequence 38, Appl
42	17	1.1	2712	4 US-08-984-709A-38	Sequence 38, Appl
43	17	1.1	2712	4 US-08-450-272-38	Sequence 38, Appl
44	17	1.1	2715	3 US-08-777-147-2	Sequence 2, Appl1
45	17	1.1	2970	3 US-08-949-386-37	Sequence 37, Appl

ALIGNMENTS

RESULT 1
US-09-327-138C-38
; Sequence 38, Application US/09327138C
; Patent No. 6541244
; GENERAL INFORMATION:
; APPLICANT: AUERHAMMER, CHRISTOPH J.
; APPLICANT: MELMED, SHLOMO
; TITLE OF INVENTION: SUPPRESSOR OF CYTOKINE SIGNALING
(SOSCS)-3 PROMOTER AND METHODS FOR ITS USE IN GENETIC THERAPY
; TITLE OF INVENTION: IN HUMANS
; FILE REFERENCE: P07 42591 (18810-803
CURRENT APPLICATION NUMBER: US/09/327,138C
; CURRENT FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 2044
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: Promoter
; LOCATION: (-2759)...(-716)
US-09-327-138C-38

Query Match 1.38; Score 20; DB 4; Length 2044;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 504 AGGGAAGGCTCTCCTAAGCA 523
Db 18 AGGGAAGGCTCTCCTAAGCA 37

RESULT 2
US-09-327-138C-2
; Sequence 2, Application US/09327138C
; Patent No. 6541244
; GENERAL INFORMATION:
; APPLICANT: AUERHAMMER, CHRISTOPH J.
; APPLICANT: MELMED, SHLOMO
; TITLE OF INVENTION: SUPPRESSOR OF CYTOKINE SIGNALING
(SOSCS)-3 PROMOTER AND METHODS FOR ITS USE IN GENETIC THERAPY
; TITLE OF INVENTION: IN HUMANS
; FILE REFERENCE: P07 42591 (18810-803
CURRENT APPLICATION NUMBER: US/09/327,138C
; CURRENT FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2863

;; TYPE: DNA
;; ORGANISM: Mus musculus
;; FEATURE:
;; NAME/KEY: promoter
;; LOCATION: (-2759)...(104)
US-09-327-138C-2

Query Match 1.3%; Score 20; DB 4; Length 2863;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 504 AGGGAAGGCTCTCCTAAGCA 523
Db 18 AGGGAAGGCTCTCCTAAGCA 37

RESULT 3
US-09-327-138C-36
; Sequence 36, Application US/09327138C
; Patent No. 6541244
; GENERAL INFORMATION:
; APPLICANT: AUERNHAMMER, CHRISTOPH J.
; APPLICANT: MELMED, SHLOMO
; TITLE OF INVENTION: SUPPRESSOR OF CYTOKINE SIGNALING
; TITLE OF INVENTION: (SOC3)-3 PROMOTER AND METHODS FOR ITS USE IN GENETIC THERAPY
; FILE REFERENCE: P07 42591 (18810-803
; CURRENT APPLICATION NUMBER: US/09/327,138C
; CURRENT FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 3646
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (0)...(0)
; OTHER INFORMATION: promoter sequence -2759 to +927 with nucleotides
US-09-327-138C-36

Query Match 1.3%; Score 20; DB 4; Length 3646;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 504 AGGGAAGGCTCTCCTAAGCA 523
Db 18 AGGGAAGGCTCTCCTAAGCA 37

RESULT 4
US-09-327-138C-35
; Sequence 35, Application US/09327138C
; Patent No. 6541244
; GENERAL INFORMATION:
; APPLICANT: AUERNHAMMER, CHRISTOPH J.
; APPLICANT: MELMED, SHLOMO
; TITLE OF INVENTION: SUPPRESSOR OF CYTOKINE SIGNALING
; TITLE OF INVENTION: (SOC3)-3 PROMOTER AND METHODS FOR ITS USE IN GENETIC THERAPY
; FILE REFERENCE: P07 42591 (18810-803
; CURRENT APPLICATION NUMBER: US/09/327,138C
; CURRENT FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 3665
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (0)...(0)

;; OTHER INFORMATION: promoter sequence -2759 to +927 with nucleotides
;; OTHER INFORMATION: -80 to -60 deleted
US-09-327-138C-35

Query Match 1.3%; Score 20; DB 4; Length 3665;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 504 AGGGAAGGCTCTCCTAAGCA 523
Db 18 AGGGAAGGCTCTCCTAAGCA 37

RESULT 5
US-09-327-138C-3
; Sequence 3, Application US/09327138C
; Patent No. 6541244
; GENERAL INFORMATION:
; APPLICANT: AUERNHAMMER, CHRISTOPH J.
; APPLICANT: MELMED, SHLOMO
; TITLE OF INVENTION: SUPPRESSOR OF CYTOKINE SIGNALING
; TITLE OF INVENTION: (SOC3)-3 PROMOTER AND METHODS FOR ITS USE IN GENETIC THERAPY
; FILE REFERENCE: P07 42591 (18810-803
; CURRENT APPLICATION NUMBER: US/09/327,138C
; CURRENT FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3686
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (-2759)...(927)
US-09-327-138C-3

Query Match 1.3%; Score 20; DB 4; Length 3686;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 504 AGGGAAGGCTCTCCTAAGCA 523
Db 18 AGGGAAGGCTCTCCTAAGCA 37

RESULT 6
US-09-327-138C-37
; Sequence 37, Application US/09327138C
; Patent No. 6541244
; GENERAL INFORMATION:
; APPLICANT: AUERNHAMMER, CHRISTOPH J.
; APPLICANT: MELMED, SHLOMO
; TITLE OF INVENTION: SUPPRESSOR OF CYTOKINE SIGNALING
; TITLE OF INVENTION: (SOC3)-3 PROMOTER AND METHODS FOR ITS USE IN GENETIC THERAPY
; FILE REFERENCE: P07 42591 (18810-803
; CURRENT APPLICATION NUMBER: US/09/327,138C
; CURRENT FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 3686
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (0)...(0)
; OTHER INFORMATION: promoter sequence -2759 to +927 with nucleotide at
; OTHER INFORMATION: position -74 mutated from T to A, nucleotide at
; OTHER INFORMATION: position -71 mutated from C to G, nucleotide at
; OTHER INFORMATION: position -69 mutated from G to C, and nucleotide
; OTHER INFORMATION: at position -66 mutated from A to T

US-09-327-138C-37

Query Match 1.3%; Score 20; DB 4; Length 3686;

Best Local Similarity 100.0%; Pred. No. 2.5;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 504 AGGGAAGGCTCTCCTTAAGCA 523
|||||

Db 18 AGGGAAGGCTCTCCTTAAGCA 37

RESULT 7

US-09-327-138C-1

; Sequence 1, Application US/09327138C

; Patent No. 6541244

; GENERAL INFORMATION:

; APPLICANT: AUBERNHAMMER, CHRISTOPH J.

; TITLE OF INVENTION: SUPPRESSOR OF CYTOKINE SIGNALING

; TITLE OF INVENTION: (SODS)-3 PROMOTER AND METHODS FOR ITS USE IN GENETIC THERAPY

; FILE REFERENCE: P07 42591 (19810-803

; CURRENT APPLICATION NUMBER: US/09/327,138C

; NUMBER OF SEQ ID NOS: 39

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 3940

; TYPE: DNA

; ORGANISM: Mus musculus

; FEATURE:

; NAME/KEY: promoter

; LOCATION: (-2907)...(1033)

; FEATURE:

; NAME/KEY: promoter

; LOCATION: (0)...(0)

; FEATURE:

; NAME/KEY: promoter

; LOCATION: (0)...(0)

; FEATURE:

; NAME/KEY: promoter

; LOCATION: (0)...(0)

; FEATURE:

; NAME/KEY: promoter

; LOCATION: (0)...(0)

; FEATURE:

; NAME/KEY: promoter

; LOCATION: (0)...(0)

; FEATURE:

; NAME/KEY: promoter

; LOCATION: (0)...(0)

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44M storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/961,527

CLASSIFICATION: 424

PRIORITY APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB340P1

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8504

INFORMATION FOR SEQ ID NO: 163:

SEQUENCE CHARACTERISTICS:

LENGTH: 8494 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-961-527-163

Query Match 1.3%; Score 19; DB 4; Length 8494;

Best Local Similarity 100.0%; Pred. No. 8.1;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 528 TCGTTTCAAGTTGACAA 546
|||||

Db 901 TCGTTTCAAGTTGACAA 919

RESULT 9

US-09-482-273-76/c

; Sequence 76, Application US/09482273

; Patent No. 6534631

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 71 Human Secreted Proteins

; FILE REFERENCE: P2030P1

; CURRENT APPLICATION NUMBER: US/09/482,273

; EARLIER FILING DATE: 2000-01-13

; EARLIER APPLICATION NUMBER: PCT/US99/15849

; EARLIER FILING DATE: 1999-07-14

; EARLIER APPLICATION NUMBER: 60/092,921

; EARLIER FILING DATE: 1998-07-15

; EARLIER APPLICATION NUMBER: 60/092,922

; EARLIER FILING DATE: 1998-07-15

; EARLIER APPLICATION NUMBER: 60/092,956

; EARLIER FILING DATE: 1998-07-15

; NUMBER OF SEQ ID NOS: 267

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 76

; LENGTH: 742

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (707)

; OTHER INFORMATION: n equals a,t,g, or c

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (724)

; OTHER INFORMATION: n equals a,t,g, or c

; FEATURE:

NAME/KEY: SITE
LOCATION: (726)
OTHER INFORMATION: n equals a,t,g, or c
US-09-482-273-76

Query Match 1.2%; Score 18; DB 4; Length 742;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 GAGAAAGAAACTTCT 111
|||||
Db 84 GAGAAAGAAACTTCT 67

RESULT 10
US-09-328-475C-321
Sequence 321, Application US/09328475C
Patent No. 6476207
GENERAL INFORMATION:
APPLICANT: Zhang, Jimmy
APPLICANT: Astel, Jon H.
APPLICANT: Carroll III, Eddie
APPLICANT: Endege, Wilson O.
APPLICANT: Ford, Donna M.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
APPLICANT: Steimann, Kathleen E.
TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT
TITLE OF INVENTION: ARE DIFFERENTIALLY REGULATED IN PROSTATE CANCER
FILE REFERENCE: 1532.002/200130.463
CURRENT APPLICATION NUMBER: US/09/328,475C
CURRENT FILING DATE: 1999-06-09
NUMBER OF SEQ ID NOS: 341
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 321
LENGTH: 781
TYPE: DNA
ORGANISM: Homo Sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(781)
OTHER INFORMATION: n = A,T,C or G
US-09-328-475C-321

Query Match 1.2%; Score 18; DB 4; Length 781;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 GAAAGGAAACTTCTT 113
|||||
Db 551 GAAAGGAAACTTCTT 568

RESULT 11
US-08-936-165A-8
Sequence 8, Application US/08936165A
Patent No. 6348582
GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Burnham, Martin
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Lonetto, Michael
APPLICANT: Nicholas, Richard
APPLICANT: Pratt, Julie
APPLICANT: Reichard, Richard
APPLICANT: Rosenberg, Martin
APPLICANT: Ward, Judith
TITLE OF INVENTION: NO. 6348582el Prokaryotic Polynucleotides,
TITLE OF INVENTION: Polypeptides and Their Uses
NUMBER OF SEQUENCES: 534
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SmithKline Beecham Corporation

STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,165A
FILING DATE: 24-SEP-1997
CLASSIFICATION: 536

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/027,032
FILING DATE: 24-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimm, Edward R.
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50549
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:

INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1172 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-08-936-165A-8

Query Match 1.2%; Score 18; DB 4; Length 1172;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1291 CTTGAAGCTGCTGACATG 1308
|||||
Db 810 CTTGAAGCTGCTGACATG 827

RESULT 12
US-08-795-475-2
Sequence 2, Application US/08795475
Patent No. 5965390
GENERAL INFORMATION:
APPLICANT: Bjvrck, Lars
APPLICANT: Sjvrck, Ulf
TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/795,475
FILING DATE: 11-FEB-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mcmasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 100084.402D1
TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 921 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
US-08-795-475-2
ORGANISM: Escherichia coli LE392/PHDL, DSM 7054

Query Match 1.1%; Score 17; DB 2; Length 921;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1144 GAGTCAAGGACATT 1160
|||||
Db 757 GAGTCAAGGACATT 773

RESULT 13
US-09-446-504-79
Sequence 79, Application US/09446504
Patent No. 6218150
GENERAL INFORMATION:
APPLICANT: UEMORI, Takashi
APPLICANT: SATO, Yoshimi
APPLICANT: FUJITA, Tomoko
APPLICANT: MIYAKE, Kazue
APPLICANT: MUKAI, Hiroyuki
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS
FILE REFERENCE: 1422-408PCT
CURRENT APPLICATION NUMBER: US/09/446, 504
CURRENT FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: PCT/JP98/02845
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: JP 9-187496
PRIOR FILING DATE: 1997-06-26
PRIOR APPLICATION NUMBER: JP 9-320692
PRIOR FILING DATE: 1997-11-27
NUMBER OF SEQ ID NOS: 92
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 79
LENGTH: 1005
TYPE: DNA
ORGANISM: Pyrococcus furiosus
US-09-446-504-79

Query Match 1.1%; Score 17; DB 3; Length 1005;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 309 TTGGCTCATATTTTAT 325
|||||
Db 5 TTGGCTCATATTTTAT 21

RESULT 14
US-09-712-266-79
Sequence 79, Application US/09712266
Patent No. 6333158
GENERAL INFORMATION:
APPLICANT: UEMORI, Takashi
APPLICANT: SATO, Yoshimi
APPLICANT: FUJITA, Tomoko
APPLICANT: MIYAKE, Kazue
APPLICANT: MUKAI, Hiroyuki
APPLICANT: ASADA, Kiyozo

APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS
FILE REFERENCE: 1422-408PCT
CURRENT APPLICATION NUMBER: US/09/712, 266
CURRENT FILING DATE: 2000-11-15
PRIOR APPLICATION NUMBER: US 09/446, 504
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: PCT/JP98/02845
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: JP 9-187496
PRIOR FILING DATE: 1997-06-26
PRIOR APPLICATION NUMBER: JP 9-320692
PRIOR FILING DATE: 1997-11-27
NUMBER OF SEQ ID NOS: 92
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 79
LENGTH: 1005
TYPE: DNA
ORGANISM: Pyrococcus furiosus
US-09-712-266-79

Query Match 1.1%; Score 17; DB 4; Length 1005;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 309 TTGGCTCATATTTTAT 325
|||||
Db 5 TTGGCTCATATTTTAT 21

RESULT 15
US-09-107-532A-2184
Sequence 2184, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107, 532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085, 598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GRC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 2184:
SEQUENCE CHARACTERISTICS:
LENGTH: 1050 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)

;
; HYPOTHETICAL: NO
;
; ANTI-SENSE: NO
;
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
;
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (B) LOCATION 1...1050
; SEQUENCE DESCRIPTION: SEQ ID NO: 2184:
; US-09-107-532A-2184

Query Match 1.1%; Score 17; DB 4; Length 1050;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1158 ATTCAGATGGAGTGAAT 1174
 |||||
Db 947 ATTCAGATGGAGTGAAT 963

Search completed: September 2, 2003, 20:33:02
Job time : 110 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2003, 13:43:09 ; Search time 455 Seconds
(without alignments)
8881.448 Million cell updates/sec

Title: US-09-872-523-6
Perfect score: 1497
Sequence: 1 atctcgaattctgaaat.....cgaaaagracactatga 1497

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- N.Geneseq_19Jun03:*
- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
 - 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
 - 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
 - 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
 - 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
 - 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
 - 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
 - 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
 - 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
 - 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
 - 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
 - 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
 - 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
 - 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
 - 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
 - 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
 - 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
 - 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
 - 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
 - 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
 - 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
 - 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
 - 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
 - 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
 - 25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1497	100.0	1497	24	ABL52497
2	1495.4	99.9	1497	24	ABL52515
3	1495.4	99.9	1497	24	ABL52516
4	1495.4	99.9	1497	24	ABL52517
5	1103	73.7	2307	24	ABL52518
6	1103	73.7	2307	24	ABL52519
7	1103	73.7	2307	24	ABL52520
8	58.8	3.9	2969	22	AA513628

9	57.4	3.8	3590	21	AA294124	Haematopoietic ste
10	55.8	3.7	2564	25	ABX08841	Angiogenesis-asso
11	55.2	3.7	2403	22	AA527361	CDNA encoding nove
12	55.2	3.7	2968	23	ABV20957	Human prostate exp
13	55.2	3.7	2968	23	ABV26806	Human prostate exp
14	55.2	3.7	3280	24	ABQ54484	Human ovarian anti
15	53.4	3.6	1035	22	AA563158	Human purified sec
16	53.4	3.6	1838	23	AA568094	DNA encoding novel
17	53.4	3.6	2081	23	AA568090	DNA encoding novel
18	53.4	3.6	2855	18	AAV00011	Human sex comb on
19	53.4	3.6	2855	18	AAV97297	Human sex comb on
20	53.4	3.6	2855	20	AAV61313	Human Scm protein
21	53.4	3.6	3255	18	AAV97299	Human sex comb on
22	53.4	3.6	3255	18	AAV00013	Human sex comb on
23	53.4	3.6	3255	20	AAV61315	Human Scm protein
24	53.4	3.6	3327	18	AAV97298	Human sex comb on
25	53.4	3.6	3327	18	AAV00012	Human sex comb on
26	53.4	3.6	3327	20	AAV61314	Human Scm protein
27	52	3.5	204	22	ABA50121	Human breast cell
28	52	3.5	204	22	ABA35090	Probe #13556 for g
29	52	3.5	204	22	AA122966	Probe #12899 for g
30	51.8	3.5	3105	22	AAH16246	Human CDNA sequenc
31	51.6	3.4	494	22	ABA44982	Human breast cell
32	51.6	3.4	494	22	ABA25176	Probe #3642 for ge
33	51.6	3.4	494	22	AA113745	Probe #3678 for ge
34	51.4	3.4	4171	23	AA592635	DNA encoding novel
35	50.6	3.4	2380	22	AA160658	Human polynucleoti
36	50.6	3.4	2555	22	AA559625	Human cell cycle a
37	50.6	3.4	2781	22	AA158732	Human polynucleoti
38	50.6	3.4	3071	24	ABA93718	Human transmembran
39	50.6	3.4	3530	25	ACC50583	Human secreted pro
40	50.6	3.4	3530	25	ABZ71320	Secreted protein-e
41	50.2	3.4	3065	18	AAV00010	Mouse sex comb on
42	50.2	3.4	3065	18	AAV00014	Mouse sex comb on
43	50.2	3.4	3065	20	AAV61316	Mouse Scm protein
44	48	3.2	418	24	ABL68929	Kidney cancer rela
45	48	3.2	4590	22	AAH24065	Yeast AOD9604-asso

ALIGNMENTS

RESULT 1	ABL52497	standard; DNA: 1497 BP.
ID	ABL52497	
XX	ABL52497:	
AC	17-JUL-2002	(first entry)
XX		
DE	Caenorhabditis elegans lin-61 nucleotide sequence	SPQ ID NO:6.
XX		
KW	Caenorhabditis elegans; lin-6; lin-56; lin-61; tumour suppressor;	
KW	cell proliferation; nematode; cancer; gene; ds.	
OS	Caenorhabditis elegans.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..1497
FT		/*tag= a
FT		/product= "LIN-61 protein"
PN	W0200194545-A2.	
PD	13-DEC-2001.	
XX		
PF	01-JUN-2001; 2001WO-US17909.	
XX		
PR	02-JUN-2000; 2000US-208802P.	
XX		
PA	(MAST) MASSACHUSETTS INST TECHNOLOGY.	
XX		
PI	Horvitz HR, Davison EM, Lu X;	

xx		WPI: 2002-401590/43.
DR	P-PDB:	ABB78650.
xx	New Caenorhabditis elegans nucleic acid involved in tumor suppressor pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for diagnosing and treating cell proliferative diseases such as cancer -	
xx	Claim 24; Page 74; 116pp: English.	
xx	The present invention describes a substantially pure nematode C. elegans (Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide (see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide (see ABB78650), where the polypeptides comprise at least 130, 110, 130 contiguous amino acids of the 386, 322 and 498 amino acid sequences given in ABB78648 to ABB78650, and modulate cell proliferation. (I) has cytostatic activity, and can be used in gene therapy. (I) is useful for modulating proliferation of a cell, and for identifying the compound that modulates cell proliferation. (I) can be used for diagnosing an animal (preferably, human) for the presence of the cell proliferation disease, or an increased chance of developing the disease, by measuring Lin-8, Lin-56 or Lin-61 nucleic acid expression in a sample obtained from the animal, by measuring Lin-8, Lin-56 or Lin-61 polypeptide in the sample using Ab or measuring amount of lin-8, lin-56, lin-61 mRNA in the sample. (I) is also useful for identifying the compound that modulates cell proliferation. (I) and its mutants are useful for understanding cell proliferative diseases including cancer, as well as in diagnosing and treating cell proliferative diseases. (II) is also useful for identifying tumour suppressors in other species such as mammals and may be used to identify therapeutic compounds. The present sequence encodes C. elegans LIN-61 from the present invention.	
SQ	Sequence 1497 BP; 497 A; 272 G; 312 C; 416 T; 0 other;	
Query Match	100.0%; Score 1497; DB 24; Length 1497;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 1497; Conservative	0; Mismatches 0; Indels 0; Gaps 0	
OY	1 ATGCTGTGAATTTCTAAAAATTTCAGACGTAACAATAAATCGCAGCAAAACTGCATAGA	60
Dd	1 ATGCTGTGAATTTCTAAAAATTTCAGACGTAACAATAAATCGCAGCAAAACTGCATAGA	60
OY	61 ACCTACTTGGGGAATCCATTTATCATCAGTTCGAAAAGAAAACTCTTCATGCCA	120
Dd	61 ACCTACTTGTGGGAATCCATTTATCATCAGTTCGAAAAGAAAACTCTTCATGCCA	120
OY	121 GTTAGAACATTCATCGTAACCTTACAGTTAAATTTTTAAAGCAATGCGTGAAGAGAGTT	180
Dd	121 GTTAGAACATTCATCGTAACCTTACAGTTAAATTTTTAAAGCAATGCGTGAAGAGAGTT	180
OY	181 ATCTTCGAAACAGTGTCCATGATTTATGACAAAGAACTGCGATTCGATTCAGTCAGATG	240
Dd	181 ATCTTCGAAACAGTGTCCATGATTTATGACAAAGAACTGCGATTCGATTCAGTCAGATG	240
OY	241 TTTCGACGAATGAAAAGTTGGCGGATPACAGATTTCTGCGTCACAGTTATCGGAGCTGAC	300
Dd	241 TTTCGACGAATGAAAAGTTGGCGGATPACAGATTTCTGCGTCACAGTTATCGGAGCTGAC	300
OY	301 AGCAATTTTGGCTCAATATTTTATTCGACAGATATGTTGGTTGGCAAACC CGCAATG	360
Dd	301 AGCAATTTTGGCTCAATATTTTATTCGACAGATATGTTGGTTGGCAAACC CGCAATG	360
OY	361 AGTGATCCC AATATGATPAAAATTGTATGCTCCGCGCTTGCAATCAACGAAGAAATAC	420
Dd	361 AGTGATCCC AATATGATPAAAATTGTATGCTCCGCGCTTGCAATCAACGAAGAAATAC	420
OY	421 CAATATGATATGATPAAAATTGTATGATATATTCGATGATGCGCAATGCTGCGCCAAC	480
Dd	421 CAATATGATATGATPAAAATTGTATGATATATTCGATGATGCGCAATGCTGCGCCAAC	480
OY	481 TGCGTGCTCCAAATTTGATGACGAAGAGCGCTCCCTTAAGACAGATCGTTTCAAAGTT	540
Dd	481 TGCGTGCTCCAAATTTGATGACGAAGAGCGCTCCCTTAAGACAGATCGTTTCAAAGTT	540

QY	541	GGACAAGCTCTGAACTTAATTAATTAATTCACAACTCTACTGAAATACGGCTGAGCGAAT	600
Db	541	GGACAAGCTCTGAACTTAATTAATTAATTCACAACTCTACTGAAATACGGCTGAGCGAAT	600
QY	601	CAAGAAATATGTGACACGACGAATGAATGATCTATCTATCTACAAAGAAAGACTTCCGAAATCG	660
Db	601	CAAGAAATATGTGACACGACGAATGAATGATCTATCTATCTACAAAGAAAGACTTCCGAAATCG	660
QY	661	CTTCCAGATGCGAATGACGACAGCAAGTCTTTAGCTCTGGATCTCAATATTTGGATAGAC	720
Db	661	CTTCCAGATGCGAATGACGACAGCAAGTCTTTAGCTCTGGATCTCAATATTTGGATAGAC	720
QY	721	GAGGAAAGCTCTTCATATTTCTCCAGTTGGATTTGACAGACGATGGAATATCACTAAAT	780
Db	721	GAGGAAAGCTCTTCATATTTCTCCAGTTGGATTTGACAGACGATGGAATATCACTAAAT	780
QY	781	GGCAAAAAAGAAATATATTGAGCGACACAAATATAAATGCTCTAAGCAATPAAAAATGAGAA	840
Db	781	GGCAAAAAAGAAATATATTGAGCGACACAAATATAAATGCTCTAAGCAATPAAAAATGAGAA	840
QY	841	AATCCAAATATGTACTGACGACGACGACATTTGATATCAATTAAGCAAAAGATCCAAATGAT	900
Db	841	AATCCAAATATGTACTGACGACGACGACATTTGATATCAATTAAGCAAAAGATCCAAATGAT	900
QY	901	CCCATGATTTGAGAAAAGTTAAGTTGGAGCAAAAAGTTTGAGCTCATCGACGCCCTTGCT	960
Db	901	CCCATGATTTGAGAAAAGTTAAGTTGGAGCAAAAAGTTTGAGCTCATCGACGCCCTTGCT	960
QY	961	CAGCAATTCATTAACCTCCACAGTGGCTTCGATTTCTCAATTTTGCAAAACCTGAAGGATAT	1020
Db	961	CAGCAATTCATTAACCTCCACAGTGGCTTCGATTTCTCAATTTTGCAAAACCTGAAGGATAT	1020
QY	1021	CTTATTTGGGAAATGAGATGATGTCGCAAGTSCATCTTGAAGACACTTTTCTATTCATATCAAT	1080
Db	1021	CTTATTTGGGAAATGAGATGATGTCGCAAGTSCATCTTGAAGACACTTTTCTATTCATATCAAT	1080
QY	1081	AATACATTATGTCTCCAGTTGGTTATGCGGAAAAATATAATTTGGAACTTGTTCGCGCA	1140
Db	1081	AATACATTATGTCTCCAGTTGGTTATGCGGAAAAATATAATTTGGAACTTGTTCGCGCA	1140
QY	1141	GATAGTTCGAAGGAACCTTCAGATGGGATGAATACCTTGGGAAAAAGATCTGCGAAAGC	1200
Db	1141	GATAGTTCGAAGGAACCTTCAGATGGGATGAATACCTTGGGAAAAAGATCTGCGAAAGC	1200
QY	1201	CTACCGCTGACTGTGTTCCAGCCCAATGCTTCCCAAGAGAGATTAGACAAATTTAAGSTA	1260
Db	1201	CTACCGCTGACTGTGTTCCAGCCCAATGCTTCCCAAGAGAGATTAGACAAATTTAAGSTA	1260
QY	1261	ATTCTGATTTCCAAACGGGTAGGACTACGCTTGAAGTCTGTGACATGTGGAAAAATCAG	1320
Db	1261	ATTCTGATTTCCAAACGGGTAGGACTACGCTTGAAGTCTGTGACATGTGGAAAAATCAG	1320
QY	1321	TTTATTTGTCAGCTACGACGTGAATCACTTCATGGAAGACGTGATTAATGTCAATTTTCGAC	1380
Db	1321	TTTATTTGTCAGCTACGACGTGAATCACTTCATGGAAGACGTGATTAATGTCAATTTTCGAC	1380
QY	1381	GGCTGGATGGAAGATTTGATGAACCTGATGATGTGAGTCCCATGATATTTCTACCGGATA	1440
Db	1381	GGCTGGATGGAAGATTTGATGAACCTGATGATGTGAGTCCCATGATATTTCTACCGGATA	1440
QY	1441	GGATGCTGTGAAGCGCACAGTTATGTTCTACACACTCCGAAAAAGATACACTATTGA	1497
Db	1441	GGATGCTGTGAAGCGCACAGTTATGTTCTACACACTCCGAAAAAGATACACTATTGA	1497
RESULT 2			
ABLS2515			
ID	ABLS2515	standard; DNA; 1497 BP.	
AC	ABLS2515;		
XX			
DT	17-JUL-2002 (first entry)		

XX	C. elegans mutant lin-61 DNA sequence lin-61(n3446) SEQ ID NO:73.
DE	
XX	Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;
KW	cell proliferation; nematode; cancer; mutant; gene; ds.
XX	
OS	Caenorhabditis elegans.
XX	
FH	Key
FT	Location/Qualifiers
FT	1..1497
FT	/*tag= a
PN	/product= "LIN-61(n3446) protein"
XX	
PD	WO200194545-A2.
XX	
PD	13-DEC-2001.
XX	
PF	01-JUN-2001; 2001WO-US17909.
XX	
PR	02-JUN-2000; 2000US-208802P.
XX	
PA	(MASI) MASSACHUSETTS INST TECHNOLOGY.
XX	
PI	Horvitz HR, Davison EM, Lu X;
XX	
DR	WPI; 2002-401590/43.
XX	
DR	P-PSDB; ABB78697.
XX	
PT	New Caenorhabditis elegans nucleic acid involved in tumor suppressor
PT	pathway; encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
PT	diagnosis and treating cell proliferative diseases such as cancer
XX	
XX	Claim 27; Page 112-113; 116pp; English.

[illegible]

QY	181	ATCTTCGAAACAGTGTGCTCATGATTATGACAAAGACTGCGATTTGGATTCAAGTCAGATG	240
Db	181	ATCTTCGAAACAGTGTGCTCATGATTATGACAAAGACTGCGATTTGGATTCAAGTCAGATG	240
QY	241	TTTGGACAAATTGAAAAAGTTTGGCGATACAGAGTGTGGCTGCAATTATGCGAGCTGAC	300
Db	241	TTTGGACAAATTGAAAAAGTTTGGCGATACAGAGTGTGGCTGCAATTATGCGAGCTGAC	300
QY	301	ACGAAATTTTGGCTCAATATTTTATCGGACGATATGTTTGGTTGGCAACGCCGCAATG	360
Db	301	ACGAAATTTTGGCTCAATATTTTATCGGACGATATGTTTGGTTGGCAACGCCGCAATG	360
QY	361	AGTGATCCCATATNGATTAATAATTGTATATGCTCCGCGCTTGGCAATCAACGAAATPAC	420
Db	361	AGTGATCCCATATNGATTAATAATTGTATATGCTCCGCGCTTGGCAATCAACGAAATPAC	420
QY	421	CAAAATGATATGTAAATTTATGTAAATTAATGCAATTGATGGCGAAATCGTCGCGCAACT	480
Db	421	CAAAATGATATGTAAATTTATGTAAATTAATGCAATTGATGGCGAAATCGTCGCGCAACT	480
QY	481	TGCGTGTCTCCAAAATTTGGATGAGGGAAGGCTCTCTAGCAAGCATCGTTTCAAAGTT	540
Db	481	TGCGTGTCTCCAAAATTTGGATGAGGGAAGGCTCTCTAGCAAGCATCGTTTCAAAGTT	540
QY	541	GGACAAAGCTGTGAACCTATPATAATTATTCGAATTCATGAAATACGGGTAGCGCGAATT	600
Db	541	GGACAAAGCTGTGAACCTATPATAATTATTCGAATTCATGAAATACGGGTAGCGCGAATT	600
QY	601	CAAGAAATATGTGACGACGAGATGAATGTATCTATCACAAGAAAGACTTCCGGAATCG	660
Db	601	CAAGAAATATGTGACGACGAGATGAATGTATCTATCACAAGAAAGACTTCCGGAATCG	660
QY	661	CTTCAGATNGAGATGAGACAGACAACAATCTTTAGCTGTGATGTCATATTTGGATGAC	720
Db	661	CTTCAGATNGAGATGAGACAGACAACAATCTTTAGCTGTGATGTCATATTTGGATGAC	720
QY	721	GAGGGAACCTCTTCATATTTTCTGTTGGATTTGCAGACAGTCAATGATATCAACTAAAT	780
Db	721	GAGGGAACCTCTTCATATTTTCTGTTGGATTTGCAGACAGTCAATGATATCAACTAAAT	780
QY	781	GGGAAAAAGGAATATATGACACACAAATAAATTTGTCGAAGCAATPAAAAATGGAGAA	840
Db	781	GGGAAAAAGGAATATATGACACACAAATAAATTTGTCGAAGCAATPAAAAATGGAGAA	840
QY	841	AATCCAAATATGACTCGACGACGCTCACAATTTGTCATTAAGCAAAAGATCCAAATGAT	900
Db	841	AATCCAAATATGACTCGACGACGCTCACAATTTGTCATTAAGCAAAAGATCCAAATGAT	900
QY	901	CCCATGATTTGGAGAAAAGTTAAGTTGGACAAAAGTTTGAAGCTATCGACCCCTTGCT	960
Db	901	CCCATGATTTGGAGAAAAGTTAAGTTGGACAAAAGTTTGAAGCTATCGACCCCTTGCT	960
QY	961	CAGCAATTCATTAACCTCCAGCTGCTCGATTCGAAATTTTGCAAAACGGAAGATPAC	1020
Db	961	CAGCAATTCATTAACCTCCAGCTGCTCGATTCGAAATTTTGCAAAACGGAAGATPAC	1020
QY	1021	CTTATTTGGGAATGAGATGATCGACAGATGCACTTGAAGCACTTTTCCATTCATATCAAT	1080
Db	1021	CTTATTTGGGAATGAGATGATCGACAGATGCACTTGAAGCACTTTTCCATTCATATCAAT	1080
QY	1081	AATACATTTATGTTCCCGATTGGTTATGCGGAAAAGTATTAATTTTGGAACTTGTTCGCCCA	1140
Db	1081	AATACATTTATGTTCCCGATTGGTTATGCGGAAAAGTATTAATTTTGGAACTTGTTCGCCCA	1140
QY	1141	GATGAGTTTCAAGGAACATTCAGATNGGATGAATACTGTGGAAGAAGATCTGCGAAGAC	1200
Db	1141	GATGAGTTTCAAGGAACATTCAGATNGGATGAATACTGTGGAAGAAGATCTGCGAAGAC	1200
QY	1201	CTACCGCTTGACTGTGTTCAAGCCCAATGCTTTCCCAAGAGAGATTAGCAAAATTTAAGTGA	1260
Db	1201	CTACCGCTTGACTGTGTTCAAGCCCAATGCTTTCCCAAGAGAGATTAGCAAAATTTAAGTGA	1260

QY 1261 ATTCTGATTTCCAAAGCGGTAGACCTACGCCCTTGAAGCTGCTGACATGTTGTGAAATCAG 1320
DB 1261 ATTTCTGATTTCCAAAGCGGTAGACCTACGCCCTTGAAGCTGCTGACATGTTGTGAAATCAG 1320
QY 1321 TTATATTGTCACAGCTCAGTGAATCAGTTTCATGGAAGCTGATTAATGTCATATTTCGAC 1380
DB 1321 TTATATTGTCACAGCTCAGTGAATCAGTTTCATGGAAGCTGATTAATGTCATATTTCGAC 1380
QY 1381 GGCTGGGATGAAGAAATTTGATGAACGTATGATGAGTCCCATGATATTTTCACGATA 1440
DB 1381 GGCTGGGATGAAGAAATTTGATGAACGTATGATGAGTCCCATGATATTTTCACGATA 1440
QY 1441 GGATGGTGTGAAGCGACAGTATGTTCTTACAACTCCGAAAAAGTACACTATTGA 1497
DB 1441 GGATGGTGTGAAGCGACAGTATGTTCTTACAACTCCGAAAAAGTACACTATTGA 1497

RESULT 3
ABL52516
ID ABL52516 standard; DNA; 1497 BP.
XX
AC ABL52516;
XX
DT 17-JUL-2002 (first entry)
XX
DE C. elegans mutant lin-61 DNA sequence lin-61(n3447) SEQ ID NO:74.
XX
XX Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;
KM cell proliferation; nematode; cancer; mutant; gene; ds.
XX
OS Caenorhabditis elegans.
XX
FH Key
FT Location/Qualifiers
FT 1..1497
FT CDS
FT /*lag- a
FT /product= "LIN-61(n3447) protein"
XX
PN W0200194545-A2.
XX
PD 13-DEC-2001.
XX
PE 01-JUN-2001; 2001MO-US17909.
XX
PR 02-JUN-2000; 2000US-208802P.
XX
PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
PI Horvitz HR, Davison EM, Lu X;
XX
XX WPI; 2002-401590/43.
DR P-PDB; ABB78698.
XX
XX
PT New Caenorhabditis elegans nucleic acid involved in tumor suppressor
PT pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
PT diagnosing and treating cell proliferative diseases such as cancer
XX
PS Claim 27, Page 113, 116pp: English.
XX
XX The present invention describes a substantially pure nematode C. elegans
CC (Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide
CC (see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide
CC (see ABB78650), where the polypeptides comprise at least 130, 110, 130
CC contiguous amino acids of the 386, 322 and 498 amino acid sequences given
CC in ABB78648 to ABB78650, and modulate cell proliferation. (I) has
CC cytostatic activity, and can be used in gene therapy. (I) is useful for
CC modulating proliferation of a cell, and for identifying the compound that
CC modulates cell proliferation. (I) can be used for diagnosing an animal
CC (preferably, human) for the presence of the cell proliferation disease,
CC or an increased chance of developing the disease, by measuring lin-8,
CC lin-56 or lin-61 nucleic acid expression in a sample obtained from the
CC animal, by measuring LIN-8, LIN-56 or LIN-61 polypeptide in the sample
CC using Ab or measuring amount of lin-8, lin-56, lin-61 mRNA in the sample.
CC (I) is also useful for identifying the compound that modulates cell

CC proliferation. (I) and its mutants are useful for understanding cell
CC proliferative diseases including cancer, as well as in diagnosing and
CC treating cell proliferative diseases. (I) is also useful for identifying
CC tumour suppressors in other species such as mammals and may be used to
CC identify therapeutic compounds. The present sequence encodes a
CC C. elegans mutant LIN-61 protein from the present invention.
XX
SQ Sequence 1497 BP; 498 A; 272 C; 311 G; 416 T; 0 other;

Query Match 99.9%; Score 1495.4; DB 24; Length 1497;
Best local similarity 99.9%; Pred. No. 0;
Matches 1496; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTGAATTTCTGAAAAATTTGTCAGACTTAACAAAAATGCGACGAAAACTCATAG 60
DB 1 ATGCTGAATTTCTGAAAAATTTGTCAGACTTAACAAAAATGCGACGAAAACTCATAG 60
QY 61 ACCTACTGTGGGAATCCATTTTACATGATCGAAGGAAAAATCTTTCTTCA 120
DB 61 ACCTACTGTGGGAATCCATTTTACATGATCGAAGGAAAAATCTTTCTTCA 120
QY 121 GTTGAAGCATTCATCGTACCTTACAGTTAATTTTAAAGATGCGTAGAGAGAGTT 180
DB 121 GTTGAAGCATTCATCGTACCTTACAGTTAATTTTAAAGATGCGTAGAGAGAGTT 180
QY 181 ATCTTCGAAACAGTGTGCTCATGATATGACAAAGACCTGATTCATCAAGTCAGATG 240
DB 181 ATCTTCGAAACAGTGTGCTCATGATATGACAAAGACCTGATTCATCAAGTCAGATG 240
QY 241 TTTGACAGCAATGAAAGATTGGCGATACAGATGTTGCGTCAGATTATCGAGACTGAC 300
DB 241 TTTGACAGCAATGAAAGATTGGCGATACAGATGTTGCGTCAGATTATCGAGACTGAC 300
QY 301 ACGAAATTTTGGCTCAATATTTTATCGACGATATGTTGTTGGCAACGCCCAATG 360
DB 301 ACGAAATTTTGGCTCAATATTTTATCGACGATATGTTGTTGGCAACGCCCAATG 360
QY 361 AGTGATCCCAATATGATATTAATTTGATATGCTCCGCCGCTTGCAATCAAGCAATATG 420
DB 361 AGTGATCCCAATATGATATTAATTTGATATGCTCCGCCGCTTGCAATCAAGCAATATG 420
QY 421 CAAAATGATATGTAATTTATGTAATTAATTTGATGATGAGGGAATGTCGCCCAACT 480
DB 421 CAAAATGATATGTAATTTATGTAATTAATTTGATGATGAGGGAATGTCGCCCAACT 480
QY 481 TCGCTGTCTCAAAATTCGATGAAGGAAAGCTCTCTAGCAGACATCGTTCAAAGTT 540
DB 481 TCGCTGTCTCAAAATTCGATGAAGGAAAGCTCTCTAGCAGACATCGTTCAAAGTT 540
QY 541 GGACAAAGCTTGAATTAATTAATTTCCATTTCTAGCAATACGGGTAGCGGAAAT 600
DB 541 GGACAAAGCTTGAATTAATTAATTTCCATTTCTAGCAATACGGGTAGCGGAAAT 600
QY 601 CAAGAAATATGTCGACGACGATGATGATGATGATGATGATGATGATGATGATGATG 660
DB 601 CAAGAAATATGTCGACGACGATGATGATGATGATGATGATGATGATGATGATGATG 660
QY 661 GTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
DB 661 GTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
QY 721 GAGGGAAGCTTCTTCAATTTCTGTTGATTTGAGAGATGATGATGATGATGATGATG 780
DB 721 GAGGGAAGCTTCTTCAATTTCTGTTGATTTGAGAGATGATGATGATGATGATGATG 780
QY 781 GCGAAAAAGATTAATTTGAGCACAATAAATTTGTCAGCAAGCAATTAATTAATTAAT 840
DB 781 GCGAAAAAGATTAATTTGAGCACAATAAATTTGTCAGCAAGCAATTAATTAATTAAT 840
QY 841 AATCCAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
DB 841 AATCCAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900


```
QY 901 CCCATGATTTGGAGAAAAGTTAAGTTGGACAAAGTTTGAGCTCATCGACCCCTTGCGT 960
DB 901 CCCATGATTTGGAGAAAAGTTAAGTTGGACAAAGTTTGAGCTCATCGACCCCTTGCGT 960
QY 961 CAGCAATTCAATTAACCTCCACGCTGCTTGATTTCTCAAAATTTTGGCAAAACCTGAGAGATAT 1020
DB 961 CAGCAATTCAATTAACCTCCACGCTGCTTGATTTCTCAAAATTTTGGCAAAACCTGAGAGATAT 1020
QY 1021 CTATATTTGGGAATGATGATGCTCCAGATGCACTTGAAGACAAATTTCTATTCATCAAT 1080
DB 1021 CTATATTTGGGAATGATGATGCTCCAGATGCACTTGAAGACAAATTTCTATTCATCAAT 1080
QY 1081 AATACATTTATGTTTCCAGTTGGTTATGGGAAAAGTATTAATTTGGAACCTTGTCCGCA 1140
DB 1081 AATACATTTATGTTTCCAGTTGGTTATGGGAAAAGTATTAATTTGGAACCTTGTCCGCA 1140
QY 1141 GATGAGTTCAAGGAACATTCAGATGATGAATCACTTGAGAGAAATCTGCAGAAACC 1200
DB 1141 GATGAGTTCAAGGAACATTCAGATGATGAATCACTTGAGAGAAATCTGCAGAAACC 1200
QY 1201 CTACCGCTTGACTGTTTCAAGCCATGCTCCGCCAGAGAGATTAGACAAATTTAAGTA 1260
DB 1201 CTACCGCTTGACTGTTTCAAGCCATGCTCCGCCAGAGAGATTAGACAAATTTAAGTA 1260
QY 1261 ATTCTGATTTCCAAACGGGTAGGACTACGCTTGAAGCTGCTGACATGTGTGAATACTAG 1320
DB 1261 ATTCTGATTTCCAAACGGGTAGGACTACGCTTGAAGCTGCTGACATGTGTGAATACTAG 1320
QY 1321 TTTATTTTGTCCAGCTTCACTGAATGATGATGAGAACTGATTAATGCTCAATTTTCGAC 1380
DB 1321 TTTATTTTGTCCAGCTTCACTGAATGATGATGAGAACTGATTAATGCTCAATTTTCGAC 1380
QY 1381 GGCTGGGATGAAGAAATTTGATGAACCTGTATGATGTGAGCTCCCATGATATTTACCGATA 1440
DB 1381 GGCTGGGATGAAGAAATTTGATGAACCTGTATGATGTGAGCTCCCATGATATTTACCGATA 1440
QY 1441 GGATGGTGTGAAGCCGACAGTTATGTTCTACAACTCTCCGAAAAAGTAACTATTTGCA 1497
DB 1441 GGATGGTGTGAAGCCGACAGTTATGTTCTACAACTCTCCGAAAAAGTAACTATTTGCA 1497

RESULT 4
ABL52517
ID ABL52517 standard; DNA; 1497 BP.
XX
AC ABL52517;
XX
DT 17-JUL-2002 (first entry)
XX
DE C. elegans mutant lin-61 DNA sequence lin-61(n3624) SEQ ID NO:75.
XX
KM Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;
XX
KW cell proliferation; nematode; cancer; mutant; gene; ds.
XX
OS Caenorhabditis elegans.
XX
FH key Location/Qualifiers
FT CDS 1..1497
FT /tag= a
FT /product= "lin-61(n3624) protein"
FT
XX
XX MO200194545-A2.
XX
XX 13-DEC-2001.
XX
XX 01-JUN-2001: 2001MO-US17909.
XX
XX 02-JUN-2000: 2000US-208802P.
XX
XX (MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX
XX Horvitz HR, Davison EM, Lu X;
```

```
DR WPI: 2002-401590/43.
DR P-SDS: ABB78699.
XX
PT New Caenorhabditis elegans nucleic acid involved in tumor suppressor
PT pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
PT diagnosing and treating cell proliferative diseases such as cancer
XX
PS Claim 27, Page 113-114; 116pp; English.
XX
CC The present invention describes a substantially pure nematode C. elegans
CC (Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide
CC (see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide
CC (see ABB78650), where the polypeptides comprise at least 130, 110, 130
CC contiguous amino acids of the 386, 322 and 498 amino acid sequences given
CC in ABB78648 to ABB78650, and modulate cell proliferation. (I) has
CC cytosolic activity, and can be used in gene therapy. (I) is useful for
CC modulating proliferation of a cell, and for identifying the compound that
CC modulates cell proliferation. (I) can be used for diagnosing an animal
CC (preferably, human) for the presence of the cell proliferation disease,
CC or an increased chance of developing the disease, by measuring lin-8,
CC lin-56 or lin-61 nucleic acid expression in a sample obtained from the
CC animal, by measuring amount of lin-8, lin-56, lin-61 mRNA in the sample.
CC using Ab or measuring amount of lin-8, lin-56, lin-61 mRNA in the sample.
CC (I) is also useful for identifying the compound that modulates cell
CC proliferation. (I) and its mutants are useful for understanding cell
CC proliferative diseases including cancer, as well as in diagnosing and
CC treating cell proliferative diseases. (I) is also useful for identifying
CC tumour suppressors in other species such as mammals and may be used to
CC identify therapeutic compounds. The present sequence encodes a
CC C. elegans mutant lin-61 protein from the present invention.
XX
SQ Sequence 1497 BP; 497 A; 271 C; 312 G; 417 T; 0 other;
XX
Query Match 99.9%; Score 1495.4; DB 24; Length 1497;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1496; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGTCTGAATTTCTGAAAATTTGTACAGAGCTTACAAAAAATCGACAGAAAATCTGATTAAG 60
DB 1 ATGTCTGAATTTCTGAAAATTTGTACAGAGCTTACAAAAAATCGACAGAAAATCTGATTAAG 60
QY 61 ACCTACTTGTGGGAATTCCTATTATTACATCACTGCTGAGAAAGAAAACCTTCTTCAATCCA 120
DB 61 ACCTACTTGTGGGAATTCCTATTATTACATCACTGCTGAGAAAGAAAACCTTCTTCAATCCA 120
QY 121 GTTGAAGCAATTCATGCTAACCTTACAGTAATTTTAAAGCAATGCGTGAAGAGAGATT 180
DB 121 GTTGAAGCAATTCATGCTAACCTTACAGTAATTTTAAAGCAATGCGTGAAGAGAGATT 180
QY 181 ATCTTGAAGAACAGTGTCCATGATTTATGACAAAGCTGCGATTCAGTCAAGTCAGATGG 240
DB 181 ATCTTGAAGAACAGTGTCCATGATTTATGACAAAGCTGCGATTCAGTCAAGTCAGATGG 240
QY 241 TTTGCAAGAAATGAAAAAGTTTGGATACAGAGTTCTGGCTCAGTTATCGAGCTGAC 300
DB 241 TTTGCAAGAAATGAAAAAGTTTGGATACAGAGTTCTGGCTCAGTTATCGAGCTGAC 300
QY 301 ACGAAATTTTGGCTCAATATTTTATGAGCAGATATGTTGGTTCGAAAGCGCGCAATG 360
DB 301 ACGAAATTTTGGCTCAATATTTTATGAGCAGATATGTTGGTTCGAAAGCGCGCAATG 360
QY 361 AGTGATCCCAATATGATTAATTTATATGCTCCGCGCTTCGAATCAAGAAAGATATAC 420
DB 361 AGTGATCCCAATATGATTAATTTATATGCTCCGCGCTTCGAATCAAGAAAGATATAC 420
QY 421 CAAAATGATATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
DB 421 CAAAATGATATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
QY 481 TCGCTGTCTCCAAATTCGATGAGGAGGAGGCTCTCTTAAGCAAGCATCTTTCAAAGTT 540
DB 481 TCGCTGTCTCCAAATTCGATGAGGAGGAGGCTCTCTTAAGCAAGCATCTTTCAAAGTT 540
```


Db	707	GTGCTCCATGATTTATGACAAAGAACTGGCATTCGATTCAGTCAGATGGTTTGACGAATT	766
Qy	253	GAAGAAAGTTTCGGATACAGAGTTCTGGCTCAGTTTATCGGAGCTGACACGAATTTTGG	312
Db	767	GAAGAAAGTTTCGGATACAGAGTTCTGGCTCAGTTTATCGGAGCTGACACGAATTTTGG	826
Qy	313	CTCAATATTTTATCGGAGCATATGTTGGTTGGCAA	350
Db	827	CTCAATATTTTATCGGAGCATATGTTGGTTGGCAAAGTAAGTTGACGCTCAGCTCTT	886
Qy	351	-----	350
Db	887	TCCTACTATTTAAATAAATAATGCTTCTGTACATAAAATCTAGAGACAATCGTATTA	946
Qy	351	-----CGCCGCAATGAGTGAT	366
Db	947	AAAGTCGAACATTGTGTATAATGATAAATTTAGACATTTACAGCGCCGCAATAGTGAT	1006
Qy	367	CCCAATATGATATAAATTTGTAATGCTCGCGCTTGCAATCAACGAAGAATACCAAAAT	426
Db	1007	CCCAATATGATATAAATTTGTAATGCTCGCGCTTGCAATCAACGAAGAATACCAAAAT	1066
Qy	427	GATATGTAATTTATGTAAT	447
Db	1067	GATATGTAATTTATGTAATGTAATGTAAGTTGTTTTTTTCCGAATTTATGTTAATATCATCT	1128
Qy	448	-----AATTGCAATGATGGCGAAATCGTCGGCAAACTTCGCTGCTCCAAAT	496
Db	1127	CACAACCTTCAGAAATTCATGATGATGGCGAAATCGTCGGCAAACTTCGCTGCTCCAAAT	1186
Qy	497	TCGATGAAAGGGAAGGCTCTCCTTAAGCAAGCATCGTTTAAAGTTTGGACACGTCTTGAC	556
Db	1187	TCGATGAAAGGGAAGGCTCTCCTTAAGCAAGCATCGTTTAAAGTTTGGACACGTCTTGAC	1248
Qy	557	TATTTAAATTTATCCAAATCTACTGTAATTAACGGTACCGCAATCAACAAATATGTGAC	616
Db	1247	TATTTAAATTTATCCAAATCTACTGTAATTAACGGTACCGCAATCAACAAATATGTGAC	1306
Qy	617	GACGAATGAATGTATCTATCAACAAAGAAAGACTTCCGAATCGCTTCAGATGCAGATG	676
Db	1307	GACGAATGAATGTATCTATCAACAAAGAAAGACTTCCGAATCGCTTCAGATGCAGATG	1366
Qy	677	ACGACAGCAAGCTCTTACGTCTGATCTCAATATGATATGACAGAGGAAGCTTCTTCA	736
Db	1367	ACGACAGCAAGCTCTTACGTCTGATCTCAATATGATATGATATGACAGAGGAAGCTTCTTCA	1426
Qy	737	TATTTCCGTTGGATTTGACAGCATGGAATATCACTAAATGCGAAAGGAATTA	796
Db	1427	TATTTCCGTTGGATTTGACAGCATGGAATATCACTAAATGCGAAAGGAATTA	1486
Qy	797	TTGAGCACACAAATAAATTTGCTCAGCAATATAAATAAGAGAAATCCAGATATGACT	856
Db	1487	TTGAGCACACAAATAAATTTGCTCAGCAATATAAATAAGAGAAATCCAGATATGACT	1546
Qy	857	CAGACAGCATCATTTGATCATTTAGCAAAAGATCCATTTGATCCATGATTTGGAGAA	916
Db	1547	CAGACAGCATCATTTGATCATTTAGCAAAAGATCCATTTGATCCATGATTTGGAGAA	1606
Qy	917	AAGTTAAGGTTGGAACAAAGTTTGAAGTCATCGACACCCTTGGCTCAGCAATTTCAATACC	976
Db	1607	AAGTTAAGGTTGGAACAAAGTTTGAAGTCATCGACACCCTTGGCTCAGCAATTTCAATACC	1666
Qy	977	TCACAGTCGCTCATTTCTCAAAATTTTSCAAACAGAGAGATTCCTTATTTGGGAATGG	1036
Db	1667	TCACAGTCGCTCATTTCTCAAAATTTTSCAAACAGAGAGATTCCTTATTTGGGAATGG	1726
Qy	1037	ATGCTCAGATGACCTTGAGACAGTTTCTCTATTCATATCATATATCATTTATGTTCC	1096
Db	1727	ATGCTCAGATGACCTTGAGACAGTTTCTCTATTCATATCATATATCATTTATGTTCC	1786
Qy	1097	CAGTTGCTTATCGGAAAGATATATTTTGGAACTTTGCCGCGAGATGATTTCAAGAA	1156
Db	1787	CAGTTGCTTATCGGAAAGATATATTTTGGAACTTTGCCGCGAGATGATTTCAAGAA	1846

QY	1157	CATTCAGATGGGATGGAATACCTTGGAGAAAGAAATTCGAGAAACCCCTACCGTTACTTGT	1216
Db	1847	CATTTCAGATGGGATGGAATACCTTGGAGAAAGAAATTCGAGAAACCCCTACCGTTACTTGT	1906
QY	1217	TCAGGCCAATGCCCTCCCAAGAGAGATTAGACCAATTTAAGCTAATTCGATTTCCAAAC	1276
Db	1907	TCAGGCCAATGCCCTCCCAAGAGAGATTAGACCAATTTAAGCTAATTCGATTTCCAAAC	1966
QY	1277	G-----	1277
Db	1967	GGTGTGTTTATATCGTTTGAGATTGTTTCACATTTAATAGTATATCATATTTGTTCTT	2026
QY	1278	-----CGTAGCACTACCGCTTGAAGCTGCTGACATGTCGAAATTCAGTTATTTGTC	1330
Db	2027	GTTTAAAGTGTGAGCTACGCTTCGATGAGCTGTCACATGTCGAAATTCAGTTATTTGTC	2086
QY	1331	CAGCTACGTAATAATCAGTTCATGAGAGACATGATAAATGTCATATTTGACGCGTGGATG	1390
Db	2087	CAGCTACGTAATAATCAGTTCATGAGAGACATGATAAATGTCATATTTGACGCGTGGATG	2146
QY	1391	AAGATTGATGAAGTGTATGATGCA-----	1418
Db	2147	AAGATTGATGAAGTGTATGATGCGAGTGAGTTATCAGACCGAACATTTTTC	2206
QY	1419	-----CTCCATGATTTCTACCGATGAGATGGTGAAGCGC	1456
Db	2207	AATGAAATTTCTATCATTTTCAGCTCCCTCATGATTTCTACCGATGAGATGGTGAAGCGC	2266
QY	1457	ACAGTTATGTTCTACCAACCTCCGAAAAAGTACCAACTATTGA	1497
Db	2267	ACAGTTATGTTCTACCAACCTCCGAAAAAGTACCAACTATTGA	2307
RESULT 6			
ABLS52519			
ID	ABLS52519	standard; DNA; 2307 BP.	
XX	AC	ABLS52519;	
XX	XX		
DT	17-JUL-2002	(first entry)	
XX	XX		
DE	C. elegans lin-61(sy223)	DNA sequence SEQ ID NO:77.	
XX	XX		
KW	Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;		
XX	XX	cell proliferation; nematode; cancer; mutant; gene; ds.	
OS	Caenorhabditis elegans.		
XX	OS		
PN	WO200194545-A2.		
PD	13-DEC-2001.		
XX	XX		
PF	01-JUN-2001; 2001WO-US17909.		
XX	PR	02-JUN-2000; 2000US-208802P.	
XX	PA	(MASI) MASSACHUSETTS INST TECHNOLOGY.	
XX	XX		
PI	Horvitz HR, Davison EM, Lu X;		
XX	XX		
DR	WPI; 2002-401590/43.		
XX	XX		
PT	New Caenorhabditis elegans nucleic acid involved in tumor suppressor		
PT	pathway, encoding LIN-8; LIN-56 or LIN-61 polypeptide, useful for		
PT	diagnosing and treating cell proliferative diseases such as cancer		
XX	PS	Disclosure; Page 115; 116pp; English.	
XX	XX		
CC	The present invention describes a substantially pure nematode C. elegans		
CC	(Caenorhabditis elegans) nucleic acid (1) encoding a LIN-8 polypeptide		
CC	(see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide		
CC	(see ABB78650), where the polypeptides comprise at least 130, 110, 130		

CC contiguous amino acids of the 386, 322 and 498 amino acid sequences given
CC in ABB578648 to ABB187630, and modulate cell proliferation. (1) has
CC cytosolic activity, and can be used in gene therapy. (1) is useful for
CC modulating proliferation of a cell, and for identifying the compound that
CC modulates cell proliferation. (1) can be used for diagnosing an animal
CC (preferably, human) for the presence of the cell proliferation disease,
CC or an increased chance of developing the disease, by measuring lln-8,
CC lln-56 or lln-61 nucleic acid expression in a sample obtained from the
CC animal, by measuring lln-8, lln-56 or lln-61 polypeptide in the sample
CC using Ab or measuring amount of lln-8, lln-56, lln-61 mRNA in the sample.
CC (1) is also useful for identifying the compound that modulates cell
CC proliferation. (1) and its mutants are useful for understanding cell
CC proliferative diseases including cancer, as well as in diagnosing and
CC treating cell proliferative diseases. (1) is also useful for identifying
CC tumour suppressors in other species such as mammals and may be used to
CC identify therapeutic compounds. The present sequence represents a
CC *C. elegans* lln-61(sy223) DNA sequence from the present invention.
XX

Sequence 2307 BP: 741 A: 393 C: 465 G: 708 T: 0 other:

Sequence 2307 BP; 741 A; 393 C; 465 G; 708 T; 0 other;

Query match 73.78; Score 1103; DB 24; Length 2307;

Best Local Similarity 81.3%; Pred. No. 2.2e-283;

Matches 1497; Conservative 0; Mismatches 0; Indels 344; Gaps 5;

OY	1	ATGTCGTAATTTCTGAAAATTTCTGAGAGCTTACACAAAATTCGACAGAAAACCTGATATAG	60
Db	467	ATGTCGTAATTTCTGAAAATTTGTGACAGCTTACACAAAATTCGACAGAAAACCTGATATAG	526
OY	61	ACCTACTGTGGGAATCCTATTATTCATCATGTTGAGAAAGGAAAACTCTTTCATTCCA	120
Db	527	ACCTACTGTGGGAATCCTATTATTCATCATGTTGAGAAAGGAAAACTCTTTCATTCCA	586
OY	121	GTGGAAGCATTCATTCGTAACTTACAGTTAAATTTTAACGAATCGTGAAGGAAGAGTT	180
Db	587	GTGGAAGCATTCATTCGTAACTTACAGTTAAATTTTAACGAATCGTGAAGGAAGAGTT	646
OY	181	ATC-----TTGGAACA	192
Db	647	ATCGAGCTTCATATTTGTCGTAAATCGTTTTAAATACAAATTTTGTAGTTCGAAACA	706
OY	193	GTGGTCATGATTTATGACAAAGACGTGCATTCGATTCAGTCAGATGTTTGCAGAAAT	252
Db	707	GTGGTCATGATTTATGACAAAGACGTGCATTCGATTCAGTCAGATGTTTGCAGAAAT	766
OY	253	GAAGAAAGTTTGCAGATACAGAGTTCTGCTCAGTTTATCGAGCTGACACGAAATTTTGG	312
Db	767	GAAGAAAGTTTGCAGATACAGAGTTCTGCTCAGTTTATCGAGCTGACACGAAATTTTGG	826
OY	313	CTCAATATTTTATCGAGACATATGTTGGTTGGCAA-----	350
Db	827	CTCAATATTTTATCGAGACATATGTTGGTTGGCAAAGTAAAGTTGACGCTCAAGCTCTT	886
OY	351	-----	350
Db	887	TCTACTATTCTTAAATTAATATGTTCTGTTACATAAATTCAGACAAACATCSTATTAT	946
OY	351	-----CGCCGCATGAGTAT	366
Db	947	AAACTTCGAACATTTGTATATACTAAATTTTGAACATTTTCAGCGCGCATGAGTAT	1006
OY	367	CCCAATATGATAAATTTGTATATGCTCCGCGCTTGCATATCAAGAAATATCCAAAT	426
Db	1007	CCCAATATGATAAATTTGTATATGCTCCGCGCTTGCATATCAAGAAATATCCAAAT	1066
OY	427	GATATGTAATTAATGATAAT-----	447
Db	1067	GATATGTAATTAATGATAATGTAATTTGTTTTTCCGAATTTATGTAATATCATCT	1128
OY	448	-----AATGCAATGATGGCGAAATCGTGCGCAAACTCGCTGCTCCAAAT	496
Db	1127	CACAACTTCAGAAATTCATGATGGCGAAATCGTGCGCAAACTCGCTGCTCCAAAT	1186
OY	497	TGATGAAGGGAAGGCTCTCCTAACGAAGCATCGTTTCAAGATTGGACACGCTTTGAAC	556

Dd	1187	TCGATGAAGGAGAGGCTCTCCTTAAGCAAGCATCGTTCCAAAGTTGGACAACGCTTGAAC	1246
Qy	557	TATTAATAATTAATCCAAATCTACTGAATAATACGCGTAGCGCGAATTCAGAAATATGTGGAC	616
Dd	1247	TATTAAATATATCCAAATCTACTGAATAATACGCGTAGCGCGAATTCAGAAATATGTGGAC	1306
Qy	617	GACGAGATGAATGTATCTATACAAAGAAAGACTTCCGAATCGCTTCCAGATGCAGATG	676
Dd	1307	GACGATGAAAGTATCTATACAAAGAAAGACTTCCGAATCGCTTCCAGATGCAGATG	1366
Qy	677	ACGACAGACAGTCGTTAGCTCTGGAATCTCAATTTGGATAGACGAGGAAAGCTTCTTCA	736
Dd	1367	ACGACAGACAAAGCTTTAGCTCTGGAATCTCAATTTGGATAGACGAGGAAAGCTTCTTCA	1426
Qy	737	TATTTCTGTTGATTTTGGACAGCTCAATGATATCAATTAATGCGAAAAAGGATATTA	796
Dd	1427	TATTTCTGTTGATTTTGGACAGCTCAATGATATCAATTAATGCGAAAAAGGATATTA	1486
Qy	797	TTTGAGACACAAATTAATAATTTGCTACAGCAATTAATAATTTGGAAAAATCCAAAGATATGCT	856
Dd	1487	TTTGAGACACAAATTAATAATTTGCTACAGCAATTAATAATTTGGAAAAATCCAAAGATATGCT	1546
Qy	857	CAGACAGCTGACATTTGATCAATTAGCAAAAGATCCAAATGATCCCATGATTTGGAGAA	916
Dd	1547	CAGACAGCTGACATTTGATCAATTAGCAAAAGATCCAAATGATCCCATGATTTGGAGAA	1606
Qy	917	AAGTTAAGGTTGGACAAAAGTTTGAAGTCATGACCCCTTGCTCAGCAATTCATTAAC	976
Dd	1607	AAGTTAAGGTTGGACAAAAGTTTGAAGTCATGACCCCTTGCTCAGCAATTCATTAAC	1666
Qy	977	TCACAGTCGCTTCATCTTCCAAATTTTGGCAAACCTGAAGATATCTTATTTGGGATATG	1036
Dd	1667	TCACAGTCGCTTCATCTTCCAAATTTTGGCAAACCTGAAGATATCTTATTTGGGATATG	1726
Qy	1037	ATGTCGCGATGACCTTGAAGACGTTTCCATTCATATCAATAATACATTAATGTATCC	1096
Dd	1727	ATGTCGCGATGACCTTGAAGACGTTTCCATTCATATCAATAATACATTAATGTATCC	1786
Qy	1097	CAGTTGGTTATGCGGAAAAGTATATTTTGGAACTGTTCCGCGCAGATGAGTTCAAAGAA	1156
Dd	1787	CAGTTGGTTATGCGGAAAAGTATATTTTGGAACTGTTCCGCGCAGATGAGTTCAAAGAA	1846
Qy	1157	CATTAGATGGGATGAATATCTTGGAGAAAGAAATTCGCAAAACCCCTACCGCTGACTGT	1216
Dd	1847	CATTAGATGGGATGAATATCTTGGAGAAAGAAATTCGCAAAACCCCTACCGCTGACTGT	1906
Qy	1217	TCAACCCATAGGCTTCCCAAGAGAGATAGACAAATTAATGAATCTGATTTCCAAAC	1276
Dd	1907	TCAACCCATAGGCTTCCCAAGAGAGATAGACAAATTAATGAATCTGATTTCCAAAC	1966
Qy	1277	G-----	1277
Dd	1967	GGGTTGTTTATATGCTTGGAGATGTTTCACTATTAATAGTTATTCATAATGTCTTCT	2026
Qy	1278	-----GCTAGACTACGCTTGAAGCTGCTGACATCTGGAATAACAGTTTATTTGTCT	1330
Dd	2027	GTTTAAAGTATAGACTACGCTTGAAGCTGCTGACATCTGGAATAACAGTTTATTTGTCT	2086
Qy	1331	CAGCTACAGTGAATAACAGTTTCAATGGAACAGTAAATGTCATAATTTGCACGCTGGGATG	1390
Dd	2087	CAGCTACAGTGAATAACAGTTTCAATGGAACAGTAAATGTCATAATTTGCACGCTGGGATG	2146
Qy	1391	AAGAATTTGATGAACCTGTATGATGG-----	1417
Dd	2147	AAGAATTTGATGAACCTGTATGATGGAGGTATGATCATGACCGAAACGACATTTTTC	2206
Qy	1418	-----ACTCCATGATATTTACCGATAGAGATGCTGTAAGCC	1456
Dd	2207	AATGAAATTTCTATATTTCAACTCCATGATATTTCTCCATATGAGATGCTGTAAGCC	2266
Qy	1457	ACAGTTATGTTTACAACTCCGAAAAAGTACAACTATTGA	1497

Db 2267 ACAGTATGTTCTACACCTCCGAAAAAGTACACTATTGA 2307

RESULT 7
ABL52520
ID ABL52520 standard; DNA; 2307 BP.
XX
XX ABL52520;
XX
XX 17-JUL-2002 (first entry)
XX
XX C. elegans lin-61(n3635) DNA sequence SEQ ID NO:78.
DE
XX
XX Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;
KM cell proliferation; nematode; cancer; mutant; gene; ds.
XX
XX Caenorhabditis elegans.
OS
XX WO200194545-A2.
PN
XX 13-DEC-2001.
PD
XX 01-JUN-2001; 2001WO-US17909.
PF
XX 02-JUN-2000; 2000US-208802P.
PR
XX (MASI) MASSACHUSETTS INST TECHNOLOGY.
PA
XX Horvitz HR, Davison EM, Lu X;
PI
XX WPI; 2002-401590/43.
DR
XX
XX New Caenorhabditis elegans nucleic acid involved in tumor suppressor
PT pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
PT diagnosing and treating cell proliferative diseases such as cancer
XX
XX Disclosure; Page 116; 116pp; English.
XX
XX The present invention describes a substantially pure nematode C. elegans
CC (Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide
CC (see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide
CC (see ABB78650), where the polypeptides comprise at least 130, 110, 130
CC contiguous amino acids of the 386, 322 and 498 amino acid sequences given
CC in ABB78648 to ABB78650, and modulate cell proliferation. (I) has
CC cytostatic activity, and can be used in gene therapy. (I) is useful for
CC modulating proliferation of a cell, and for identifying the compound that
CC modulates cell proliferation. (I) can be used for diagnosing an animal
CC (preferably, human) for the presence of the cell proliferation disease,
CC or an increased chance of developing the disease, by measuring lin-8,
CC lin-56 or lin-61 nucleic acid expression in a sample obtained from the
CC animal, by measuring LIN-8, LIN-56 or LIN-61 polypeptide in the sample.
CC using Ab or measuring amount of lin-8, lin-56, lin-61 mRNA in the sample.
CC (I) is also useful for identifying the compound that modulates cell
CC proliferation. (I) and its mutants are useful for understanding cell
CC proliferative diseases including cancer, as well as in diagnosing and
CC treating cell proliferative diseases. (I) is also useful for identifying
CC tumour suppressors in other species such as mammals and may be used to
CC identify therapeutic compounds. The present sequence represents a
CC C. elegans lin-61(n3635) DNA sequence from the present invention.
XX
SQ Sequence 2307 BP; 741 A; 393 C; 465 G; 708 T; 0 other;

Query Match 73.7%; Score 1103; DB 24; Length 2307;
Best Local Similarity 81.3%; Pred. No. 2,2e-283;
Matches 1497; Conservative 0; Mismatches 0; Indels 344; Gaps 5;

QY 1 AAGTCTGAATTTCTGAAAAATGTCAGAGCTAACAAAAATCGACAGAAAACTCGATAG 60
DB 467 AAGTCTGAATTTCTGAAAAATGTCAGAGCTAACAAAAATCGACAGAAAACTCGATAG 526

QY 61 ACCTACTTGTGGGAATCCTATTATTCATCGATCGATTCGAGAAAGAAAACTTCTTCATTC 120
DB 527 ACCTACTTGTGGGAATCCTATTATTCATCGATCGATTCGAGAAAGAAAACTTCTTCATTC 586

QY 121 GTTGAAGCATTCATCGTAACCTTACAGTTAATTTTAAAGAAATCGTGAAGAGAGATT 180
DB 587 GTTGAAGCATTCATCGTAACCTTACAGTTAATTTTAAAGAAATCGTGAAGAGAGATT 646

QY 181 ATC-----TTCGAAGA 192
DB 647 ATCGTAGTTTCATATTGTTTCGTAATCGTTTTAAAAACAATTTTGTAGTTCGAAGA 706

QY 193 GTGTCATGATTTATATGACAAAGAACTCGATTCGATTCAGTCAAGTGTTCGACAAATT 252
DB 707 GTGTCATGATTTATGACAAAGAACTCGATTCGATTCAGTCAAGTGTTCGACAAATT 766

QY 253 GAAAAATTGTCGATACAGAGTTCGTCGATTCATTCGAGACGTACAGAAATTTTGG 312
DB 767 GAAAAATTGTCGATACAGAGTTCGTCGATTCATTCGAGACGTACAGAAATTTTGG 826

QY 313 CTCGAATATTTTATCGACGATATGTTTGGTTCGCAAA----- 350
DB 827 CTCGAATATTTTATCGACGATATGTTTGGCAAAAGTAAGTTGAGCCTCAGCTTT 886

QY 351 ----- 350
DB 887 TCTACTATCTAAATAAATAATGTTCTGTACATAAAATTCTAGAGAAACATCGATTTA 946

QY 351 -----CGCCGCAATGAGTAT 366
DB 947 AAACCTCGAAACATTTGTATATAGTAAATTTGAACATTTTCAGCCGCCCAATGATGAT 1006

QY 367 CCCAATATGATATAAATTTGTATATGCTCCGCCCTTCGAATCAAGCAAGAAATACAAAT 426
DB 1007 CCCAATATGATATAAATTTGTATATGCTCCGCCCTTCGAATCAAGCAAGAAATACAAAT 1066

QY 427 GATATGTAATATATGTAAT----- 447
DB 1067 GATATGTAATATATGTAATATGTAATGTTGTTTTTCCGAATTTATGTTAATATCATCT 1126

QY 448 -----AATTCATGATGCGCAAAATCGCGCCAACTCGCTGTCGCAAAAT 496
DB 1127 CACAACTTCGAAATTTGATGATGCGCAAAATCGCGCCAACTCGCTGTCGCAAAAT 1186

QY 497 TCGATGAAGGAAGGCTCTCTTAAGCAAGCATGTTTCAAAAGTTGGAACAAGTCTTGAC 556
DB 1187 TCGATGAAGGAAGGCTCTCTTAAGCAAGCATGTTTCAAAAGTTGGAACAAGTCTTGAC 1246

QY 557 TATTAATATATTCATATTCATGTAATACGCTACCGCAATTCAGAAATATGTGAGAC 616
DB 1247 TATTAATATATTCATATTCATGTAATACGCTACCGCAATTCAGAAATATGTGAGAC 1306

QY 617 GACGATGAATGATCTATCATCAAAAGAAAGACTTTCGCCAATGCTTCCAGATGACAGAT 676
DB 1307 GACGATGAATGATCTATCATCAAAAGAAAGACTTTCGCCAATGCTTCCAGATGACAGAT 1366

QY 677 ACGACAGACAGCTTTAGCTGTGATCTCAATATTTGGATAGACGAGGAAAGCTTCTCA 736
DB 1367 ACGACAGACAGCTTTAGCTGTGATCTCAATATTTGGATAGACGAGGAAAGCTTCTCA 1426

QY 737 TATTTCTGTGATTTGACGAGTCGATGATATCACTAATTTGGAAAAAGCAATATA 796
DB 1427 TATTTCTGTGATTTGACGAGTCGATGATATCACTAATTTGGAAAAAGCAATATA 1486

QY 797 TTGACGACAAATTAATAATTTGCTCAAGCAATAAATAATGAGAAATTCGAATATGACT 856
DB 1487 TTGACGACAAATTAATAATTTGCTCAAGCAATAAATAATGAGAAATTCGAATATGACT 1546

QY 857 CACAGCAGCTCACATTTGATCAATATAGCAAAAGATCCATTTGATCCATGATTTGGAGAA 916
DB 1547 CACAGCAGCTCACATTTGATCAATATAGCAAAAGATCCATTTGATCCATGATTTGGAGAA 1606

QY 917 AAGTTAAGTTGGACAAAAGTTTGAAGTATGACATGACCCCTTGGCTCAGCAATTAATPACC 976
DB 1607 AAGTTAAGTTGGACAAAAGTTTGAAGTATGACATGACCCCTTGGCTCAGCAATTAATPACC 1666

```

QY 977 TCACAGTCGCTTCGATTCGCAAAATTTGCAAAAAGTGAAGATCTTATTTGGGGANGG 1036
DB 1667 TCACAGTCGCTTCGATTCGCAAAATTTGCAAAAAGTGAAGATCTTATTTGGGGANGG 1726
QY 1037 ATGGTCAGATGACATGTAAGACAGTTTCCTATTCATATCAATATATACATTTATGTTCC 1096
DB 1727 ATGGTCAGATGACATGTAAGACAGTTTCCTATTCATATCAATATATACATTTATGTTCC 1786
QY 1097 CAGTTGGTTATGCGGAAAAAGTATTAATTTGAACTGTTCCGCGAGATGATTCMAAGNA 1156
DB 1787 CAGTTGGTTATGCGGAAAAAGTATTAATTTGAACTGTTCCGCGAGATGATTCMAAGNA 1846
QY 1157 CATTCGATGCGATGTAATCTTGAGAGAAAGATCTGCAAAACCCTACCGCTTGACTTGT 1216
DB 1847 CATTCGATGCGATGTAATCTTGAGAGAAAGATCTGCAAAACCCTACCGCTTGACTTGT 1906
QY 1217 TCAGCCAAATGCTTCCCAAGAGAGATTAAGCAAAATTAAGTAAATTCGATTTCCAAAC 1276
DB 1907 TCAGCCAAATGCTTCCCAAGAGAGATTAAGCAAAATTAAGTAAATTCGATTTCCAAAC 1966
QY 1277 G----- 1277
DB 1967 GGGTGTGTTTATATCGTTGAGATGTTTCACATTAATAGTTATTCATATTTGTTCTT 2026
QY 1278 -----GGTAGACACGCTTGAGCTGACATGTTGTAAGTCAATTCATTTATTTGTC 1330
DB 2027 GTTTTAAGTAGACACGCTTGAGCTGACATGTTGTAAGTCAATTCATTTATTTGTC 2086
QY 1331 CAGCTACAGTGAATGATTCATGAGAGCTGATTAATTCATTTGCGAGCTGGGATG 1390
DB 2087 CAGCTACAGTGAATGATTCATGAGAGCTGATTAATTCATTTGCGAGCTGGGATG 2146
QY 1391 AGCAATTTGATGAACGTATGATGGA----- 1418
DB 2147 AGCAATTTGATGAACGTATGATGAGTGAAGTTATTCATGACCGAAGACATTTTTC 2206
QY 1419 -----CTCCATGATTTCTACCGATAGGATGGTGAAGCGC 1456
DB 2207 AATGAAATTTCTATTCATTTCAAGCTCCCATGATATTTTACCGATAGGATGGTGAAGCGC 2266
QY 1457 ACAGTTATGTTCTACAACTCCGAAAAAGTACAACTATTGA 1497
DB 2267 ACAGTTATGTTCTACAACTCCGAAAAAGTACAACTATTGA 2307

RESULT 8
AAS13628 standard; cDNA: 2969 BP.
XX AAS13628:
XX AAS13628:
XX 18-DEC-2001 (first entry)
XX cDNA sequence encoding mammalian Spg14.
DE cDNA sequence encoding mammalian Spg14.
XX Mammalian; reproductive-specific protein; male infertility; gene therapy;
KW spermatogenesis; sperm count disorder; anti infertility;
KW reproduction; ss.
XX Mammalia.
OS Mammalia.
XX M0200166752-A2.
XX 13-SEP-2001.
XX 07-MAR-2001; 2001MO-US07371.
XX 07-MAR-2000; 2000US-0187518.
XX 12-JAN-2001; 2001US-0261557.
XX (WHEED ) WHITEHEAD INST BIOMEDICAL RES.
PA Wang PJ, Page DC;
PI

```

```

XX WPI: 2001-570774/64.
DR P-PSDB: AAU07864.
XX Novel reproduction-specific protein, useful for treating disorders of
PT reduced sperm count, enhancing/increasing sperm count and/or sperm
PI activity.
XX Claim 1; Fig 11; 151pp; English.
XX The present invention relates to the isolation of novel mammalian and
PS human reproductive-specific proteins (AAU07859-AAU07899), and the
CC nucleic acids encoding them. The nucleic acids encoding
CC reproductive-specific proteins are useful for diagnosing infertility
CC which is a result of reduced sperm count, reduced sperm motility,
CC malformed sperm or combinations of these. The sequences of the invention
CC are useful as markers for spermatogenic cells, for identifying genes or
CC proteins characteristic of male infertility, diagnosing or aiding in
CC the diagnosis of infertility in men, and for contraception in which
CC sperm production or sperm count is reduced or defective sperm is
CC produced. Antibodies to reproductive-specific proteins are useful for
CC determining the presence of these proteins in a sample obtained from a
CC man being assessed for infertility, for identifying the expression of
CC genes in particular cell type or particular developmental stage, for
CC studies of spermatogenesis, and for immunofluorescence of germ cells or
CC in Western blots for assessing the presence of the protein the antibody
CC binds. The sequences of the invention are also useful for treating
CC disorders of reduced sperm count, and for increasing sperm count and/or
CC sperm activity. The nucleic acids of the invention are useful in gene
CC therapy. AAS13623-AAS13647 represent cDNA sequences encoding for the
CC mammalian reproduction-specific proteins of the present invention.
XX Sequence 2969 BP; 923 A; 629 G; 689 G; 728 T; 0 other:

Query Match 3.9%; Score 58.8; DB 22; Length 2969;
Best Local Similarity 55.3%; Pred. No. 2.9e-05;
Matches 114; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 1279 GTAGACATACGCTTGAAGCTGCTGACATGCTGTGAATAATCACTTATTTGCCAGCTCA 1338
DB 801 GTTGGAGATGAAGATGAGCTGTGCTGATGAAAAAATTCATTTATGATGCTGCCACA 860
QY 1339 GTGAATGATGATGAGAGAGCTGATTAATGCTCAATTTGCAAGCTGGATGAAGATTT 1398
DB 861 ATTGAGAGCTGTCTGTGAGATCACTTATCATATCACTTTTGATGATGAGTGAAGCATTT 920
QY 1399 GATGAACGTATGATGAGTGGACTCCCATGATATTTTACCGATAGATGGTGTGAAGCGCAC 1458
DB 921 GATTAATGTTGTGACTATGACTCCGAGACATCTTCCAGTTGATGGTGTGCCCTCACA 980
QY 1459 AGTTATGTTCTACAACTCCGAAAA 1484
DB 981 GGAGATGATTTACACCGACACAGAAA 1006

RESULT 9
AA294124 standard; cDNA: 3590 BP.
XX AA294124:
XX AA294124:
XX 19-JUN-2000 (first entry)
XX Haematopoietic stem cell specific nucleic acid.
DE Haematopoietic stem cell specific nucleic acid.
XX Haematopoietic stem cell; immune system disorder;
KW leukaemia; antileukaemic; immunomodulator; therapy; mouse; ss.
XX Mus musculus.
OS Mus musculus.
XX Key Location/Qualifiers
FH 260..2155
FT CDS /*tag= a

```

```

FT      /transl_except= (pos:1748..1750, aa:Val)
FT      /transl_except= (pos:1826..1828, aa:Arg)
FT      /transl_except= (pos:1901..1903, aa:Asn)
PN      WO200011168-A2.
PD      02-MAR-2000.
XX      20-AUG-1999; 99WO-US19052.
XX      21-AUG-1998; 98US-0138132.
XX      (UYPR-) UNIV PRINCETON.
XX      Lemischka I, Moore K;
XX      WPI: 2000-237650/20.
XX      P-PSDB; AAY79183.
DR      Hematopoietic stem cell signaling proteins modulating replication and
PT      differentiation for treating immune system disorders and leukemia -
PS      Claim 10; Page 228-229; 256pp; English.

XX      The present sequence is that of a nucleic acid isolated from
XX      mouse primitive stem cells by a method of the invention. It is an
XX      example of claimed isolated nucleic acids (see AYZ94077-294131) that
XX      are specifically expressed in haematopoietic stem cells (HSCs) and
XX      which encode HSC-specific proteins. The HSCs are especially
XX      primitive HSCs (PHSCs) such as umbilical cord cells, bone marrow
XX      cells and foetal liver cells. The encoded proteins (see AAY79176-93)
XX      are growth factors, transcription factors, splicing factors, capping
XX      factors, transport proteins, translation factors or replication
XX      factors that modulate HSC activity, especially differentiation or
XX      replication. The invention provides a claimed method for
XX      identifying PHSC-specific nucleic acids involving: creating a PHSC
XX      cDNA library and a non-PHSC immune cell library; and subtracting
XX      the 2 libraries. Also claimed are methods: for generating a stem
XX      cell/progenitor cell from PHSCs; for identifying the presence of a
XX      PHSC in a sample; for identifying the presence in a sample of a
XX      compound that modulates HSC activity; for using such a compound to
XX      treat an immune system condition, especially leukaemia; for
XX      introducing exogenous nucleic acid into a HSC; and for ex vivo
XX      expansion of HSCs. Also claimed is a PHSC specifically expressing 1
XX      of the claimed nucleic acids, such as the present sequence.

SQ      Sequence 3590 BP; 1101 A; 702 C; 797 G; 990 T; 0 other:
Query Match      3.8%; Score 57.4; DB 21; Length 3590;
Best Local Similarity 55.2%; Pred. No. 7.4e-05;
Matches 112; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

OY      1276 CGGGTAGGAGTACCGCTGAGCTGTCGATGTCGAAATATGATTATTTGTCACGCT 1335
DB      1745 CGGTTGGAGATGAAATTTAGAGCTGTAGATCTCATGAGCAGCCAGTTAAATGTGAGCC 1804
OY      1336 ACAGTAAATCAAGTTCATGAGAAAGATGATTAATGTCATTTCCAGCGCTGGATGAAGA 1395
DB      1805 ACAGTACCTCGAATTTATTCACCATCTCTTGAGGATACATTTTGATGTGGAGAAAGAG 1864
OY      1396 TTGTGATGAACGTATGATGTGACGACGCCATGATATTTCTACCGATGAGATGCTGAAGCG 1455
DB      1865 TATGACCACTGGGTAGACTGTGAGTCCCTGACCTTATCCGTAGGAGGTGTCAAGTTA 1924
OY      1456 CACAGTTATGTTCTACACCTCC 1478
DB      1925 ACTGATATCACTACAGCCTCC 1947

```

RESULT 10
 ABX08841
 ID ABX08841 standard; cDNA: 2564 BP.
 XX

```

AC      ABX08841;
XX      21-JAN-2003 (first entry)
XX      Angiogenesis-associated human polynucleotide sequence #103.
DE      Human; angiogenesis-associated transcript; angiogenesis;
XX      angiogenesis-associated disease; cancer; cytostatic; gene therapy;
XX      gene; ss.
XX      Homo sapiens.
XX      WO200279492-A2.
XX      10-OCT-2002.
XX      14-FEB-2002; 2002WO-US04915.
XX      14-FEB-2001; 2001US-0784356.
XX      22-FEB-2001; 2001US-0791390.
XX      19-APR-2001; 2001US-285475P.
XX      03-AUG-2001; 2001US-310025P.
XX      13-NOV-2001; 2001US-350666P.
XX      29-NOV-2001; 2001US-334244P.
XX      (EOSB-) EOS BIOTECHNOLOGY INC.
XX      Murray R, Glynn R, Watson SR, Aziz N;
XX      WPI: 2003-040681/03.
XX      P-PSDB; ABU03557.
DR      Detecting angiogenesis-associated transcript in a cell for diagnosing
PT      and treating cancer by contacting a sample with a polynucleotide that
PT      exhibits changes in expression level as a function of time in tissue
PT      undergoing angiogenesis.
XX      Example 2; Page 274-275; 291pp; English.
XX      The present invention relates to methods and compositions for
XX      detecting an angiogenesis-associated transcript in a cell in
XX      a patient. The method involves contacting a biological sample from
XX      the patient with a polynucleotide that selectively hybridises to a
XX      sequence at least 80% identical to any of the angiogenesis-associated
XX      human polynucleotide sequences given in the specification. These
XX      angiogenesis-associated polynucleotide sequences comprise genes that
XX      exhibit changes in expression levels as a function of time in tissue
XX      undergoing angiogenesis. The method and the polynucleotide sequences
XX      of the invention are useful for diagnosing and treating angiogenesis
XX      and angiogenesis-associated diseases e.g. cancer. The polynucleotide
XX      sequences are also useful in the gene therapy of such disorders. The
XX      angiogenesis-associated proteins encoded by the polynucleotide
XX      sequences are useful as a vaccine for therapeutic and prophylactic
XX      immunisation. ABX08739-ABX08853 represent angiogenesis-associated
XX      polynucleotide sequences.

SQ      Sequence 2564 BP; 832 A; 482 C; 560 G; 690 T; 0 other:
Query Match      3.7%; Score 55.8; DB 25; Length 2564;
Best Local Similarity 54.7%; Pred. No. 0.00017;
Matches 111; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

OY      1276 CGGGTAGGAGTACCGCTGAGCTGTCGATGTCGAAATATGATTATTTGTCACGCT 1335
DB      1153 CGGTGAGATGAAATTTAGAGCAGTAGATCTCATGAGCAGCTTTAATATCTGAGCC 1212
OY      1336 ACAGTAAATCAAGTTCATGAGAAAGATGATTAATGTCATTTCCAGCGCTGGATGAAGA 1395
DB      1213 ACAGTAACTCGAATTTATTCATCTGCTTGAGGATACATTTTGATGAGGAGAAAGAG 1272
OY      1396 TTGTGATGAACGTATGATGTGACGACGCCATGATATTTCTACCGATGAGATGCTGAAGCG 1455
DB      1273 TATGATCACTGGGTAGACTGTGAGTCACTGACCTTATCCGTAGGAGGTGTCAAGTTA 1332

```

OY 1456 CACAGTTATGTTCTACACCTCC 1478
| | | | | | | | | |
Db 1333 ACTGATATCACTACACGCTCC 1355

RESULT 11
AAS27361
ID AAS27361 standard; cDNA: 2403 BP.
XX
XX AAS27361;
XX
XX 07-NOV-2001 (first entry)
XX
XX cDNA encoding novel signal transduction pathway protein, Seq ID 396.
XX
XX Neuroprotective; cytoskeletal; dermatological; immunosuppressive; tumour;
XX antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
XX immune system disorder; rheumatoid arthritis; hepatitis C; blood disorder;
XX organ transplant rejection; infection; hepatitis; Gaucher's disease;
XX sickle cell anaemia; hyperproliferative disorder; Parkinson's disease;
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
XX cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
XX reproductive system; gastrointestinal; liver disorder; AIDS; ss;
XX acquired immune deficiency syndrome.
XX
XX Homo sapiens.
XX
XX MO200154733-A1.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01312.
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205515.
XX 07-JUN-2000; 2000US-0209467.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
XX 07-JUL-2000; 2000US-0216647.
XX 07-JUL-2000; 2000US-0216880.
XX 11-JUL-2000; 2000US-0217487.
XX 11-JUL-2000; 2000US-0217496.
XX 14-JUL-2000; 2000US-0218290.
XX 26-JUL-2000; 2000US-0220963.
XX 26-JUL-2000; 2000US-0220964.
XX 14-AUG-2000; 2000US-0224518.
XX 14-AUG-2000; 2000US-0224519.
XX 14-AUG-2000; 2000US-0225213.
XX 14-AUG-2000; 2000US-0225214.
XX 14-AUG-2000; 2000US-0225266.
XX 14-AUG-2000; 2000US-0225267.
XX 14-AUG-2000; 2000US-0225268.
XX 14-AUG-2000; 2000US-0225270.
XX 14-AUG-2000; 2000US-0225447.
XX 14-AUG-2000; 2000US-0225757.
XX 14-AUG-2000; 2000US-0225758.
XX 14-AUG-2000; 2000US-0225759.
XX 18-AUG-2000; 2000US-0226279.
XX 22-AUG-2000; 2000US-0226681.
XX 22-AUG-2000; 2000US-0226868.
XX 23-AUG-2000; 2000US-0227182.
XX 30-AUG-2000; 2000US-0228924.
XX 01-SEP-2000; 2000US-0229287.
XX 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231245.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0232968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239933.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.

PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 06-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
X1 Rosen CA, Barnash SC, Ruben SM;
XX WPI; 2001-465460/50.
DR P-PSDB; AAU17444.
XX

PT Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders -
XX
XX

PS Claim 1; SEQ ID No 396; 880bp; English.

XX
CC The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC abnormalities (Down syndrome), ischemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorders), liver disorders (cirrhosis),
CC as stimulators of B-cell responsiveness to pathogens, activators of
CC T-cells, to induce higher affinity antibodies, and as a means to induce
CC tumour proliferation in pathologies e.g. acquired immune deficiency
CC syndrome (AIDS). AAS267857-AAS27850 represent novel signal transduction
CC pathway protein coding sequences and PCR primers of the invention.
XX

Query Match	3.74;	Score 55.2;	DB 22;	Length 2403;
Best Local Similarity	55.98;	Pred. No. 0.00024;		
Matches 105; Conservative	0;	Mismatches 83;	Indels 0;	Gaps 0;

[illegible]

Db 760 GATGATCAACCGGGTAAAGTTCACCTTBTATGGCTGGAAACAATGTGATGATTACTGGATA 819

Qy 1411 GATGTGACATCCCAATGATATCTACCGATAGAGAGTGTGAAGCCGACAGTTATGTCTA 1470

Db 820 GATGACAGTTCTCTGATTAATCACTCTAGGCTGGTGTCAAAAACAGGACATCCCTT 879

Qy 1471 CAACCTTC 1478

Db 880 CAGCCTCC 887

RESULT 12
ABV20957/c
ID ABV20957 standard; cDNA; 2968 BP.

DT	13-SEP-2002 (first entry)
XX	
DE	Human prostate expression marker CDNA 20948

KW	Human: prostate cancer; cytostatic; carcinogen; pharmacodynamic marker
KM	pharmacogenomic marker; gene; ss.
XX	
OS	Homo sapiens.

PN W0200160860-A2

PD	23-AUG-2001.
XX	
PF	20-FEB-2001; 2001WO-US05171.

PR 17-FEB-2000; 2000US-183319P;
PR 16-MAR-2000; 2000US-189842P;
PR 25-MAY-2000; 2000US-207454P;
PR 09-JUN-2000; 2000US-211314P;
PR 18-JUL-2000; 2000US-219007P;
PR 13-DEC-2000; 2000US-255281P;

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Monahan JE,

DR WPI; 2001-662795/76.

PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 XX
 PS Claim 1; Page 3444; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV000010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 2968 BP, 850 A; 655 C; 555 G; 905 T; 3 other.

Query Match	3.7%	Score 55.2	DB 23	Length 2968
Best Local Similarity	55.9%	Pred. No. 0.00027		
Matches 105; Conservative	0;	Mismatches 83;	Indels 0;	Gaps 0

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

Db 508 TTGGGAGGTTGCGGGCTCAGAGTGCTTGTGCACTTTGATGGGTGGCGAGGGCCCTTGG 567
QY 1400 ATGAACGTATGATGTGAGCTCCAGATATTTCTACCGATAGATGTGTGAAGCGCACA 1459
Db 568 ACTACGTGTCGGCTTCGACACTCCGAGACATCTTCCCTGTGGGCTGTCTTCTTGACAG 627
QY 1460 GTTATGTTCTACACCTCC 1478
Db 628 GAGACAACCTGCAGCCTCC 646

RESULT 2

US-08-852-153-5
; Sequence 5, Application US/08852153
; Patent No. 5914266
; GENERAL INFORMATION:
; APPLICANT: Randazzo, Filippo
; TITLE OF INVENTION: Mammalian Sex Comb on Midleg Acts as a Tumor Suppressor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,153
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Guth, Joseph H.
; REGISTRATION NUMBER: 31,261
; REFERENCE/DOCKET NUMBER: 1224.006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 923-3888
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ. ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3235 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-852-153-5

Query Match 3.6%; Score 53.4; DB 2; Length 3235;
Best Local Similarity 54.3%; Pred. No. 1.9e-05;
Matches 108; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 1280 TAGACTACCCCTTGAAGCTGTCAGATGTGTAATAATCATTTTATTTGCCAGCTACAG 1339
Db 848 TTGGGAATGAGCTAGAGCTGTGGACAGAGAACCCTCATTTTCATTGCTCCACTCA 907
QY 1340 TGAATCAGTTTCAGAGCTGATAAATGTCATTTCGAGGCTGGGATGAGAAATTTG 1399
Db 908 TTGGGGAGGTTTGGGGCTCAGAGTCTGTGTCACTTTTGATGGGTGGCGAGGGCCCTTGG 967
QY 1400 ATGAACGTATGATGTGAGCTCCAGATATTTCTACCGATAGATGTGTGAAGCGCACA 1459
Db 968 ACTACGTGTCGGCTTCGACACTCCGAGACATCTTCCCTGTGGGCTGTCTTCTTGACAG 1027
QY 1460 GTTATGTTCTACACCTCC 1478
Db 1028 GAGACAACCTGCAGCCTCC 1046

RESULT 3

US-08-852-153-3
; Sequence 3, Application US/08852153
; Patent No. 5914266
; GENERAL INFORMATION:
; APPLICANT: Randazzo, Filippo
; TITLE OF INVENTION: Mammalian Sex Comb on Midleg Acts as a Tumor Suppressor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,153
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Guth, Joseph H.
; REGISTRATION NUMBER: 31,261
; REFERENCE/DOCKET NUMBER: 1224.006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 923-3888
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ. ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3327 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-852-153-3

Query Match 3.6%; Score 53.4; DB 2; Length 3327;
Best Local Similarity 54.3%; Pred. No. 1.9e-05;
Matches 108; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 1280 TAGACTACCCCTTGAAGCTGTCAGATGTGTAATAATCATTTTATTTGCCAGCTACAG 1339
Db 920 TTGGGAATGAGCTAGAGCTGTGGACAGAGAACCCTCATTTTCATTGCTCCACTCA 979
QY 1340 TGAATCAGTTTCAGAGCTGATAAATGTCATTTCGAGGCTGGGATGAGAAATTTG 1399
Db 980 TTGGGAGGTTTGGGGCTCAGAGTCTGTGTCACTTTTGATGGGTGGCGAGGGCCCTTGG 1039
QY 1400 ATGAACGTATGATGTGAGCTCCAGATATTTCTACCGATAGATGTGTGAAGCGCACA 1459
Db 1040 ACTACGTGTCGGCTTCGACACTCCGAGACATCTTCCCTGTGGGCTGTCTTCTTGACAG 1099
QY 1460 GTTATGTTCTACACCTCC 1478
Db 1100 GAGACAACCTGCAGCCTCC 1118

RESULT 4

US-09-620-312D-622
; Sequence 622, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom

APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungling
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: PL-Genes Version 1.0
SEQ ID NO 622
LENGTH: 2781
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (263)..(1750)
US-09-620-312D-622

Query Match 3.4%; Score 50.6; DB 4; Length 2781;
Best Local Similarity 53.2%; Pred. No. 0.00011;
Matches 107; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 1278 GGTAGAGCTACGCTTGAAGCTGCTGACATGTGTGAATAATCAATTTATTTGTCCAGCTAC 1337
DB 1243 GGTGGGCTAGTAAGCTGGAGCCGTGGACCTGATGAGACCCCGCTCATCTGTGTGGCCAC 1302
QY 1338 AGTGAATCAGTTCATGAGAGACTGATTAATGTCAATTTGACGGCTGGGATGAGAATT 1397
DB 1303 GGTGAAGAGAGTGTGATGAGCTCCACACATCCACTTTGACGGCTGGGAGACGAGCTA 1362
QY 1398 TGATGACATGTATGATGTGATGATGCCATGATTAATTCACGATAGGATGTGTGAAGCGCA 1457
DB 1363 CGACCGAGTGGTGGTACTGACGATGCCAGACATCTACCCCGCTGGGTGTGTAGCTCAC 1422
QY 1458 CAGTTATGTCTCTACAACCTCC 1478
DB 1423 CGGCTACGAGCTCCAGCTCC 1443

RESULT 5
US-08-852-153-7
Sequence 7, Application US/08852153
Patent No. 5914266
GENERAL INFORMATION:
APPLICANT: Randazzo, Filippo
TITLE OF INVENTION: Mammalian Sex Comb on Midleg Acts as a Tumor Suppressor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: California
COUNTRY: U.S.A.
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/852,153
FILING DATE:

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Guth, Joseph H.
REGISTRATION NUMBER: 31,261
REFERENCE/DOCKET NUMBER: 1224,006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 923-3888
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3065 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-852-153-7

Query Match 3.4%; Score 50.2; DB 2; Length 3065;
Best Local Similarity 53.3%; Pred. No. 0.00016;
Matches 106; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 1280 TAGACTACGCCCTTGAAGCTGCTGACATGTGTGAATAATCAATTTATTTGTCCAGCTACAG 1339
DB 864 TGGGAATGAAGTTAGAGCTGTAGACAGAAAGAACCCCTCATTTCTATTTGCCAGCCACTA 923
QY 1340 TGAATCATCTTCAATGAGAGCTGATTAATGTCAATTTGACGGCTGGGATGAGAATTTG 1399
DB 924 TTGGAGAAGATTTCGAGCGCGCAAGATGCTAGTACCTTTGATGGGTGGCGAGCGCATTTG 983
QY 1400 ATGAATCTGATGATGTGACTCCCATGATTAATTCACGATAGATGTGTGAAGCGACA 1459
DB 984 ACTACTGTGCTCCGCTTTGACTCCCGGAGACATCTTCTGTGGCTGTGTCTTTGACTG 1043
QY 1460 GTATATGTTTACAACCTCC 1478
DB 1044 GAGATACCTGCAACCCAC 1062

RESULT 6
US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU

```
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-F15
US-08-232-463-14
```

```
Query Match 2.8%; Score 41.6; DB 1; Length 7218;
Best Local Similarity 2.8%; Pred. No. 0.068;
Matches 11; Conservative 213; Mismatches 162; Indels 0; Gaps 0;
```

```
QY 580 GAATACGCTAGCGCAATTCAAGAAATATGTGACGACGAATGATATCTATCACA 639
D 1429 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1370
QY 640 AAGAAAGACTTTCGCCAGATCTCCAGATGACAGATGACAGCAAGCTTAGCTCT 699
D 1369 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1310
QY 700 GGATCTCAATATGATAGACGAGGAGGAGCTTCTCATATTTCTGTTGATTGCACA 759
D 1309 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1250
QY 760 GTCAATGATATCACTAAATGCGAAAGAATATATTAGCACACAATAAATGCT 819
D 1249 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1190
QY 820 CAAGCAATAAATAATGAGAAATCCAGATATGACTGACGACGACATTTGATCAA 879
D 1189 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1130
QY 880 TTAGCAAGAATCCATTTGATCCCATGATTGGAGAAAGTTAAGTTGACAAAGTTT 939
D 1129 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1070
QY 940 GAGCTCATGACCCCTGGCTGACGA 965
D 1069 RRRATCGAAGCTCCCTGACCTGCA 1044
```

RESULT 7

```
US-09-134-001C-212
Sequence 212, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 212
LENGTH: 1284
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-212
```

```
Query Match 2.8%; Score 41.2; DB 4; Length 1284;
Best Local Similarity 47.3%; Pred. No. 0.04;
Matches 124; Conservative 0; Mismatches 138; Indels 0; Gaps 0;
```

```
QY 677 ACACAGACAACTCTTTACTCTGATCTCAATATTGGATACAGGAGAACTCTTCA 736
D 833 ATGAAAGACAAATCAGATTAATTTAAATTAATGATGACGGTATGAAATCTGCA 892
QY 737 TATTCCCTTGATTTGACGAGTCATGATATCAATTAATCGAAAAAGAAATATA 796
D 893 ATTGGACAGGTGATAGAACAGTAATTAAGAACCAACATTAATGTTATGTTGCTC 952
QY 797 TTGAGCACACAATAAATTTGCTCAAGCAATTAATAAATGAGAAATCCAGATATGACT 856
D 953 ATAAATATGAGATATTGATGCGTTAGTACACAATAAGACATTAATATGAGAGATA 1012
QY 857 CAGACAGCTCACTTTGATTCATTAAGCAAAAGATCCATTTATCCCATTTGAGAGA 916
D 1013 AATGTATATTTTATTCTGCAATTAAGAAAGCAATTCATGATGATTAACAAT 1072
QY 917 AAGTTAAGTTGACAAAGTT 938
D 1073 TAAATGATATTGCATCAAAATT 1094
```

RESULT 8

```
US-09-134-001C-1557
Sequence 1557, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 1557
LENGTH: 1890
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1557
```

```
Query Match 2.7%; Score 40.4; DB 4; Length 1890;
Best Local Similarity 54.8%; Pred. No. 0.082;
Matches 80; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
```

```
QY 698 CTGATCTCAATTTGATAGACGAGGAGCTTCTCATATTTCTGTTGATTTGCAG 757
D 1589 CGGAATGCGAAGATGCTTGAAGATGGAACCTTTATCAAAATTAATGTTGCTTTCTA 1648
QY 758 CAGTCAATGATATCACTAAATGCGAAAAAGGATATATTGACACACAATAAATTG 817
D 1649 GAAATCTGATTAATAAAGGTATGTCACATTAATTTGTAATAATAGTGAACAATTTA 1708
QY 818 CTCAGCAATAAATAATGAGAAAT 843
D 1709 ATCGATGATTTGAATAATGGCGCTACT 1734
```

RESULT 9

```
US-08-916-421B-1
Sequence 1, Application US/08916421B
Patent No. 6503729
GENERAL INFORMATION:
APPLICANT: Built et al.
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methano
FILE REFERENCE: PB275
CURRENT APPLICATION NUMBER: US/08/916,421B
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: US 60/024,428
```



```

: PRIOR FILING DATE: 1996-08-22
: NUMBER OF SEQ ID NOS: 3
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 1
: LENGTH: 1664976
: TYPE: DNA
: ORGANISM: Methanococcus jannaschii
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (28222)..(28222)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (28257)..(28258)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (84773)..(84773)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (84808)..(84808)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (84812)..(84812)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (98120)..(98120)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (98159)..(98159)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (98239)..(98239)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (98266)..(98266)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (98343)..(98343)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (10398)..(10398)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (148948)..(148948)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (163385)..(163385)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (191989)..(191989)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (191995)..(191995)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (231980)..(231980)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (234187)..(234187)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (234220)..(234220)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (234814)..(234814)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (309398)..(309398)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (309418)..(309418)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (312837)..(312837)
: OTHER INFORMATION: n equals a, t, c, or g
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (312993)..(312993)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (319226)..(319226)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (559167)..(559167)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (559241)..(559241)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (600992)..(600992)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (622708)..(622708)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (657081)..(657081)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (657203)..(657203)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (674435)..(674435)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (682442)..(682442)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (713652)..(713652)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (741684)..(741684)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (779455)..(779455)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (779676)..(779676)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (855539)..(855539)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (871619)..(871619)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (1084830)..(1084830)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (1096846)..(1096846)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (1119881)..(1119881)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (1130881)..(1130881)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (1310988)..(1310988)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (1313224)..(1313224)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (1349473)..(1349473)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (1349491)..(1349491)
: OTHER INFORMATION: n equals a, t, c, or g
```

```

? NAME/KEY: misc.feature
? LOCATION: (1470091)..(1470091)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc.feature
? LOCATION: (1569020)..(1569020)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc.feature
? LOCATION: (1602912)..(1602912)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc.feature
? LOCATION: (1603734)..(1603734)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc.feature
? LOCATION: (1637998)..(1637998)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc.feature
? LOCATION: (1664854)..(1664855)
? OTHER INFORMATION: n equals a, t, c, or g

```

Query Match	2.68;	Score 39;	DB 4;	Length 1664976;
Best Local Similarity	46.08;	Pred. No. 4.4;		
Matches 132; Conservative	0;	Mismatches 155;	Indels 0;	Gaps 0;

OY	48	AAAAATCATATAGACCCCTGTTGTGGAAATCCATTTATCTACATCGAGTTCGAAAGAAAGAAAC	107
Db	107526	AAAAACGTAAAAACCTACTTATTTCTTCATATCATTTATATCTTGCCAAACTACAAAGAGATC	107585
OY	108	TTCTTTTCATTCACAGTTGAAGCATTCATATCGTAACCTTACACGTTATTTTAAAGCATTCGGT	167
Db	107586	TACCAATATATGATATGTGCACATTTTGAAGTTCGGTAAAAAGTGTATCTACAAAGAGAGT	107645
OY	168	GAGGAGAGAGATTATCTTCGAAACAGTGGTCCATGATTTATGACAAACAGCTGCATTTCGAT	227
Db	107646	AACCTAAAGAGATTAAAGTTAAATTTACGGTCTATGATGATGAGTAGATTTTCTAAATTTT	107705
OY	228	TCAACTCAGATGCGTTTGCAAGCAATTTGAAAAAGCTTTGCGATACAGATTCGTGCTCAGTT	287
Db	107706	TCTTTTATTTATATATACATCCATTAATTCAAATTTAAATATAAAAATCTCTTTTCCAATT	107765
OY	288	TATCGAGCTGCACAGCAATTTTGGCTCAATTTTATATCGGAGATAT	334
Db	107766	CTGAAAAAATATAAATTTTATGAAAAATATATATTTTTTTTTCATATAAAA	107812

```

1      RESULT 10
2      US-09-016-434-237
3      ; Sequence 237, Application US/09016434
4      ; Patent No. 6500938
5      ; GENERAL INFORMATION:
6      ; APPLICANT: Janice Au-Young
7      ; APPLICANT: Jeffrey J. Sellhammer
8      ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
9      ; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
10     ; NUMBER OF SEQUENCES: 1490
11     ; CORRESPONDENCE ADDRESS:
12     ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
13     ; STREET: 3174 PORTER DRIVE
14     ; CITY: PALO ALTO
15     ; STATE: CALIFORNIA
16     ; COUNTRY: USA
17     ; ZIP: 94304
18     ;
19     ; COMPUTER READABLE FORM:
20     ; MEDIUM TYPE: Floppy disk
21     ; COMPUTER: IBM PC compatible
22     ; OPERATING SYSTEM: PC-DOS/MS-DOS
23     ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
24     ; CURRENT APPLICATION DATA:
25     ; APPLICATION NUMBER: US/09/016,434
26     ; FILING DATE: HERewith
27     ; CLASSIFICATION:
28     ; PRIOR APPLICATION DATA:
29     ; APPLICATION NUMBER:
30     ;

```

```

1      FILING DATE:
2      CLASSIFICATION:
3      ATTORNEY/AGENT INFORMATION:
4      NAME: Zeller, Karen J.
5      REGISTRATION NUMBER: 37,071
6      REFERENCE/DOCKET NUMBER: PA-0002 US
7      TELECOMMUNICATION INFORMATION:
8      TELEPHONE: (650) 855-0555
9      TELEFAX: (650) 845-4166
10     INFORMATION FOR SEQ ID NO: 237:
11     SEQUENCE CHARACTERISTICS:
12     LENGTH: 216 base pairs
13     TYPE: nucleic acid
14     STRANDEDNESS: single
15     TOPOLOGY: linear
16     IMMEDIATE SOURCE:
17     LIBRARY: LINDNOT03
18     CLONE: 1577179
19     US-09-016-434-237

```

Query Match	2.6%;	Score 38.6;	DB 4;	length 216;
Best Local Similarity	58.1%;	Pred. No. 0.1;		
Matches 68;	Conservative 0;	Mismatches 49;	Indels 0;	Gaps 0;

QY	1367	ATATTCACATTTGAGGCGGTGGGATTAAGAGATTTATGGAACATCTATGAGTGGACATCCCAAG	1426
Db	5	ATATTCACATTTGAGGCGGTGGAGTGGACATTTTGTATTTACTGGTGCACAGATATGATTTCTCGAG	64
QY	1427	ATATTCACCGATATGAGTATGGTGTAAGGGCACATTTATTTCTTCAACATCTCCGAAA	1483
Db	65	ATATTTTTCACAGCTGGGTGGTGTGTGCGCTGCACAGAGATGATTTTACAAACCCCCACAGAA	121

```

RESULT 11
US-09-601-198-138/c
: Sequence 138. Application US/09601198
: Patent No. 6531583
: GENERAL INFORMATION:
: APPLICANT: Cassell, Gail H.
: APPLICANT: Chen, Ellison Y.
: APPLICANT: Glass, Jennifer S.
: APPLICANT: Glass, John I.
: APPLICANT: Helmer, Cheryl R.
: APPLICANT: Leikowitz, Elliot
: TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
: TITLE OF INVENTION: UREALYTICUM
: FILE REFERENCE: UAB-13452/22
: CURRENT APPLICATION NUMBER: US/09/601,198
: CURRENT FILING DATE: 2000-12-08
: PRIOR APPLICATION NUMBER: 60/073,189
: PRIOR FILING DATE: 1998-01-30
: NUMBER OF SEQ ID NOS: 181
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 138
: LENGTH: 747
: TYPE: DNA
: ORGANISM: Ureaplasma urealyticum
US-09-601-198-138

```

	Query Match	Similarity	2.6%	Score 38.2	DB 4	Length 747
	Best Local	Similarity	58.3%	Pred. No. 0.23		
	Matches	67	Conservative	0	Mismatches	48
					Indels	0
					Gaps	0
QY	786	AAAGAAATATATTGAGCACACAAATAAATTCCTCAACCAATTAATAATGGAGAAATCC	845			
DB	607	AAACAAATTTGCTAATGATTAACAAATACATTAATGATGATGTTTAAAAAAAAGTAAAAATCAA	548			
QY	846	AAGATATGACTCAGACAGCTCACATTTTGATCAATTAACAAAAGATCCAAATTGAT	900			
DB	547	TAAAGCGGATCAATAATTAACCTACTACTAGTTATGATTTATCAAGAAATCCAAAGAT	493			

```
US-09-601-198-127/c
; Sequence 127, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
; APPLICANT: Cassell, Gail H.
; APPLICANT: Chen, Ellison Y.
; APPLICANT: Glass, Jennifer S.
; APPLICANT: Glass, John I.
; APPLICANT: Helner, Cheryl R.
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
; FILE REFERENCE: UAB-13452/22
; CURRENT APPLICATION NUMBER: US/09/601,198
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/073,189
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 127
; LENGTH: 972
; TYPE: DNA
; ORGANISM: Ureaplasma urealyticum
US-09-601-198-127

Query Match          2.5%; Score 37.2; DB 4; Length 972;
Best Locality 54.3%; Pred. No. 0.5;
Matches 75; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 790 GAATATATTGACGACCAATTAATTTGCTCAAGCATTAATAAAGAGAGAAATCCAGA 849
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 798 GAATTTATTAGCTTTTAACTAATTAATTCATTTTGGCACCAATTAAGATATTAA 739
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 850 TATGACTGACGACGACATTCATTTGATCATTAAGCAAAAGATCCATTCATCCATGATT 909
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 738 TATGATTTATGATGACGACGATGGAATGATGATGATGATGATGATGATGATGAT 679
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 910 TGGAGAAAGTTAAGGT 927
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 678 CGTTATCATTTTAAAGGT 661
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
US-09-205-258-243
; Sequence 243, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
```

```
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 243
; LENGTH: 2271
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (553)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (2267)
; OTHER INFORMATION: n equals a,t,g, or c
```

```

? FEATURE:
? NAME/KEY: SITE
? LOCATION: (12269)
? OTHER INFORMATION: n equals a,t,g, or c
? FEATURE:
? NAME/KEY: SITE
? LOCATION: (12711)
? OTHER INFORMATION: n equals a,t,g, or c
JS-09-205-258-243
```

Query Match 2.5%; Score 37; DB 4; Length 2271;
Best Local Similarity 56.3%; Pred. No. 0.84;
Matches 67; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY	Db	QY	Db
733	2152	793	2212
TTCAATTTCTCGTTGGATTTGCAGCAATCAGATATCTCACTTAATGCGGAAAAAGGAA	TTAATATTTTGAATAGATTTTAAATAAATGCAAGATATTAATCTCAAAAAA	TATATTTGAGCACACAAATAAATTCCTCAGCAATAAATAATGCGAATAATCCAGATA	AA
851	2270		

```

: RESULT 14
: US-09-328-352-2141
: Sequence 2141, Application US/09328352
: Patent No. 6562958
: GENERAL INFORMATION:
: APPLICANT: Gary L. Breton et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
: TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: GTC99-03PA
: CURRENT APPLICATION NUMBER: US/09/328,352
: CURRENT FILING DATE: 1999-06-04
: NUMBER OF SEQ ID NOS: 8252
: SEQ ID NO 2141
: LENGTH: 879
: TYPE: DNA
: ORGANISM: Acinetobacter baumannii
: US-09-328-352-2141

```

Query Match	2.48;	Score 36.4;	DB 4;	Length 879;
Best Local Similarity	56.8%;	Pred. No. 0.81;		
Matches 67;	Conservative 0;	Mismatches 51;	Indels 0;	Gaps 0

OY	753	TGCAGCATCATGGATATTCACACTAAATGCCGAAAAAGCAATTTTGTGGACACACAATAA	812
Db	750	TCCGCGAGGAATGTGGTTGATGTCACCAACATCGAATTTCTATATTTGATCACCACCAATTA	809
OY	813	AATTGCTCAGCAATAAAAATGGAGAAATTCAGATTTGACTAGACGACGTACA	870
Db	810	ACCTGCACAAACACGGATTAAATGGAATTCACCCCAAAAATGCTACAAAATGAGGGGACA	867

```

RESULT 15
US-09-134-001C-2060
: Sequence 2060. Application US/09134001C
: Patent No. 6380370
: GENERAL INFORMATION:
: APPLICANT: Lynn Doucette-Stamm et al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
: TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: GTC-007
: CURRENT APPLICATION NUMBER: US/09/134,001C
: CURRENT FILING DATE: 1998-08-13
: PRIOR APPLICATION NUMBER: US 60/064,964
: PRIOR FILING DATE: 1997-11-08
: PRIOR APPLICATION NUMBER: US 60/055,779
: PRIOR FILING DATE: 1997-08-14
: NUMBER OF SEQ ID NOS: 5674
: SEQ ID NO 2060
:
: LENGTH: 354
: TYPE: DNA
:

```

ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2060

Query Match	2.48;	Score	36.2;	DB	4;	Length	354;
Best Local Similarity	65.48;	Pred. No.	0.62;				
Matches	53;	Conservative	0;	Mismatches	28;	Indels	0;
						Gaps	0;

Oy	783	GAAAAAAGGAATATTGACACCAAAATAAATTCCTCACAACATAAAAAATGGAGAAA	842
Db	147	GACTAAAAAGAATTTAGATGAGTAATAAAATTCCTCAATGAAAAAGGTAGAAA	206
Oy	843	TCCAAATATGACTCGACGA	863
Db	207	AGTTATAGATAACTCACATGA	227

```
Search completed: September 2, 2003, 17:43:44
Job time : 115 secs
```



```
|||||
Db 181 ATCTCGAAACAGGTCATGATTAATGACAAAGCTCGATTCGATTCAGATGAGTGG 240
Oy 241 TTTCAGCAATTTGAAAAAGTTGGGATACAGAGTTCGGCTCAAGTTATATGGAGCTAC 300
Db 241 TTTCAGCAATTTGAAAAAGTTGGGATACAGAGTTCGGCTCAAGTTATATGGAGCTAC 300
Oy 301 ACGAAATTTTGGCTCAATATTTTATCGGACGATATGTTTGGTTTGGCAACGCCGCAATG 360
Db 301 ACGAAATTTTGGCTCAATATTTTATCGGACGATATGTTTGGTTTGGCAACGCCGCAATG 360
Oy 361 AGTATCCCATATATGATTAATATGATATGCTCCGCCGCTTGCATTAAGAGAAATAC 420
Db 361 AGTATCCCATATATGATTAATATGATATGCTCCGCCGCTTGCATTAAGAGAAATAC 420
Oy 421 CAAATGATATGTAATTTATTAATTAATTCATTCATTCATTCATTCATTCATTCATTCAT 480
Db 421 CAAATGATATGTAATTTATGTAATTAATTAATTCATTCATTCATTCATTCATTCATTCAT 480
Oy 481 TCGCTGTCGCAAAATTCGATGAAGGAGGCTCTCTTACGACATCGTTTCAAAAGTT 540
Db 481 TCGCTGTCGCAAAATTCGATGAAGGAGGCTCTCTTACGACATCGTTTCAAAAGTT 540
Oy 541 GGACAACTCTTGAACCTTAATTTATTCGAATTCCTACTGAAATACGGGTAGCGGAAT 600
Db 541 GGACAACTCTTGAACCTTAATTTATTCGAATTCCTACTGAAATACGGGTAGCGGAAT 600
Oy 601 CAAGAAATATGTCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 660
Db 601 CAAGAAATATGTCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 660
Oy 661 CTTCAGATGTCAGATGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 720
Db 661 CTTCAGATGTCAGATGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 720
Oy 721 GAGGAACTCTTCATTCATTCCTGTTGATTTGAGAGTTCATTCATTCATTCATTCATTCAT 780
Db 721 GAGGAACTCTTCATTCATTCCTGTTGATTTGAGAGTTCATTCATTCATTCATTCATTCAT 780
Oy 781 GCGAAAAAGAAATATATGAGACACACAATTAATTTGCTCAAGCAATTAATTAATTAAT 840
Db 781 GCGAAAAAGAAATATATGAGACACACAATTAATTTGCTCAAGCAATTAATTAATTAAT 840
Oy 841 AATCCAGATATGACTCGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 900
Db 841 AATCCAGATATGACTCGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 900
Oy 901 CCCATGATTTGAGAAAGTTAAGGTTGAGCAAAAAGTTGAGCTCATGACCCCTTGCT 960
Db 901 CCCATGATTTGAGAAAGTTAAGGTTGAGCAAAAAGTTGAGCTCATGACCCCTTGCT 960
Oy 961 CAGCAATTCATTAACCTTCACGCTGCTTCGATTCCTCAAAATTTGCAAACTGAGAGAT 1020
Db 961 CAGCAATTCATTAACCTTCACGCTGCTTCGATTCCTCAAAATTTGCAAACTGAGAGAT 1020
Oy 1021 CTATATGTCGAAATGATGTCACAGATGCTGAAAGCACTTTTCTATTCATTCATTCAT 1080
Db 1021 CTATATGTCGAAATGATGTCACAGATGCTGAAAGCACTTTTCTATTCATTCATTCAT 1080
Oy 1081 AATCATTTATGTTCCGAGTTGTTATGCGAAAAAGTTAATTTGGAAGCTTTCGCGCA 1140
Db 1081 AATCATTTATGTTCCGAGTTGTTATGCGAAAAAGTTAATTTGGAAGCTTTCGCGCA 1140
Oy 1141 GATGAGTTCAAGAACTTCAGATGAGATTAATTCCTGGAAGAAAGATTCGCAAAACC 1200
Db 1141 GATGAGTTCAAGAACTTCAGATGAGATTAATTCCTGGAAGAAAGATTCGCAAAACC 1200
Oy 1201 CTACCGCTTGAATGTTCAAGCAATGCTTCCCAAGAGATTAAGCAAAATTTAAGSTA 1260
Db 1201 CTACCGCTTGAATGTTCAAGCAATGCTTCCCAAGAGATTAAGCAAAATTTAAGSTA 1260
Oy 1261 ATTTGATTTCCAAACGGGTAGGACTAGCCTTGAAGCTGCTGACATGTGGAATAATAG 1320
|||||
```

```
Db 1261 ATTTGATTTTCCAAACGGGTAGGACTAGCCTTGAAGCTGCTGACATGTGGAATAATAG 1320
Oy 1321 TTATTTGTCAGCTACAGTAAATCAATTCATGGAAGACTGATAAATATGCAATTTGCAC 1380
Db 1321 TTATTTGTCAGCTACAGTAAATCAATTCATGGAAGACTGATAAATATGCAATTTGCAC 1380
Oy 1381 GCGTGGATGGAAGATTTGATGAACCTGATGATGTCGACCTCCATGATATTTCAACGATA 1440
Db 1381 GCGTGGATGGAAGATTTGATGAACCTGATGATGTCGACCTCCATGATATTTCAACGATA 1440
Oy 1441 GGATGCTGTGAAGGCGACAGTATGTTCTACAACTCCGAAAAAGTCAACTATTGA 1497
Db 1441 GGATGCTGTGAAGGCGACAGTATGTTCTACAACTCCGAAAAAGTCAACTATTGA 1497
```

RESULT 2

```
US-09-872-523-73
; Sequence 73, Application US/09872523
; Patent No. US20020137906A1
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Davison, Ewa M.
; APPLICANT: Lu, Xiaowei
; TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
; FILE REFERENCE: 01997/536002
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/208,802
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 1497
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-872-523-73
```

```
Query Match 99.9%; Score 1495.4; DB 10; Length 1497;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1496; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Oy 1 ANGTCTGAATTTTGGAAAAATGTCAGAGCTTAACAAAAATGCGACAGAAAATCGATTAAG 60
Db 1 ANGTCTGAATTTTGGAAAAATGTCAGAGCTTAACAAAAATGCGACAGAAAATCGATTAAG 60
Oy 61 ACCTACTTGTGGGAATCCTATTACATCAGTTTCGAAAGGAAAAAATCTTTTCATTCCA 120
Db 61 ACCTACTTGTGGGAATCCTATTACATCAGTTTCGAAAGGAAAAAATCTTTTCATTCCA 120
Oy 121 GTTGAGCATTCATCGTAACCTTAACAGTTAATTTTAACGAATGCGTGAAGAGAGATT 180
Db 121 GTTGAGCATTCATCGTAACCTTAACAGTTAATTTTAACGAATGCGTGAAGAGAGATT 180
Oy 181 ATCTTCGAAACAGGTCATGATTAATGACAAGAACTCGATTCGATTCATTCATTCATTCAT 240
Db 181 ATCTTCGAAACAGGTCATGATTAATGACAAGAACTCGATTCGATTCATTCATTCATTCAT 240
Oy 241 TTTCGACGAATTTGAAAAAGTTTGGGATACAGAGTTCGCTCAAGTTTATGAGAGCTGAC 300
Db 241 TTTCGACGAATTTGAAAAAGTTTGGGATACAGAGTTCGCTCAAGTTTATGAGAGCTGAC 300
Oy 301 ACGAAATTTTGGCTCAATATTTTATCGGACGATATGTTTGGTTGGCAACGCCGCAATG 360
Db 301 ACGAAATTTTGGCTCAATATTTTATCGGACGATATGTTTGGTTGGCAACGCCGCAATG 360
Oy 361 AGTATCCCATATATGATTAATTAATTCGACGATATGTTTGGTTGGCAACGCCGCAATG 420
Db 361 AGTATCCCATATATGATTAATTAATTCGACGATATGTTTGGTTGGCAACGCCGCAATG 420
Oy 421 CAAATGATATGTAATTTATGTAATTAATTCGATTCGATTCGATTCGATTCGATTCGATTC 480
Db 421 CAAATGATATGTAATTTATGTAATTAATTCGATTCGATTCGATTCGATTCGATTCGATTC 480
```


OY	781	GCAGAAAAAGCATATTATTGAGCACACAAATTAATTTGCTCAGACAAATPAAAAAATGAGAA	840
Db	781	GGCAAAAAAGCAATATTATTGAGCACACAAATTAATTTGCTCAGACAAATPAAAAAATGAGAA	840
OY	841	AATCCAAGATATTGACTCAGACGACGTCCATTTGATCAATTAGCAAAAAAGATCCAAATGAT	900
Db	841	AATCCAAGATATTGACTCAGACGACGTCCATTTGATCAATTAGCAAAAAAGATCCAAATGAT	900
OY	901	CCCATGATTTGGAGAAAAAGTTAAGTTGGACAAAAGTTTGAGCTCATCGACCCCTTGCT	960
Db	901	CCCATGATTTGGAGAAAAAGTTAAGTTGGAGCAAAAAGTTTGAGCTCATCGACCCCTTGCT	960
OY	961	CAGCAATTTCAATTAACCTCCACGTGCGTTCCGATTCCTCAAAATTTGCCAAAACGAAGATAT	1020
Db	961	CAGCAATTTCAATTAACCTCCACGTGCGTTCCGATTCCTCAAAATTTGCCAAAACGAAGATAT	1020
OY	1021	CTTATTGTGGAGAAATGATGATCGATGATGCATTTGTAAGACAGTTTTCCTATTCAATCAAT	1080
Db	1021	CTTATTGTGGAGAAATGATGATCGATGATGCATTTGTAAGACAGTTTTCCTATTCAATCAAT	1080
OY	1081	AATACATTTATGTTTCCAGTTGGTTATGCGGAAAAAGTAAATTTTGAAACTTTGTTCCGCCA	1140
Db	1081	AATACATTTATGTTTCCAGTTGGTTATGCGGAAAAAGTAAATTTTGAAACTTTGTTCCGCCA	1140
OY	1141	GATAGATTTCAAGGAACATTCAGATGGGATGAATACTTGGAGAAAGAAATCGCAGAAACC	1200
Db	1141	GATAGATTTCAAGGAACATTCAGATGGGATGAATACTTGGAGAAAGAAATCGCAGAAACC	1200
OY	1201	CTACCGCTTGACTGTTTCCAGCCCAATGCTTCCACAGAGATTAAGCAATTTTAAGGTA	1260
Db	1201	CTACCGCTTGACTGTTTCCAGCCCAATGCTTCCACAGAGATTAAGCAATTTTAAGGTA	1260
OY	1261	ATTCTGATTTCCAAACGGGTGAGACTACGCTTGAAGCTGCTGACATGTGGAATAATCAG	1320
Db	1261	ATTCTGATTTCCAAACGGGTGAGACTACGCTTGAAGCTGCTGACATGTGGAATAATCAG	1320
OY	1321	TTTTTTTGTCCAGTACAGTGAATTCAGTTGATGGAAGACTGATTAATTTGCAATTTGGAC	1380
Db	1321	TTTTTTTGTCCAGTACAGTGAATTCAGTTGATGGAAGACTGATTAATTTGCAATTTGGAC	1380
OY	1381	GGCTGGAGTGAAGAATTTGATGAACCTGATGATGTGACCTCCCATGATATTCTACCGATA	1440
Db	1381	GGCTGGAGTGAAGAATTTGATGAACCTGATGATGTGACCTCCCATGATATTCTACCGATA	1440
OY	1441	GGATGCGTGAAGGCGACACTTAATGTTCTAACAACCTCCGAAAAGTACAACTATTGA	1497
Db	1441	GGATGCGTGAAGGCGACACTTAATGTTCTAACAACCTCCGAAAAGTACAACTATTGA	1497
RESULT 4			
US-09-872-523-75			
; Sequence 75, Application US/09872523			
; Patent No. US20020137906A1			
; GENERAL INFORMATION:			
; APPLICANT: Horvitz, H. Robert			
; APPLICANT: Davidson, Ewa M.			
; APPLICANT: Lu, Xiaowei			
; TITLE OR INVENTION: A Tumor Suppressor Pathway in C. Elegans			
; FILE REFERENCE: 01997/536002			
; CURRENT APPLICATION NUMBER: US/09/872,523			
; CURRENT FILING DATE: 2001-06-01			
; PRIOR APPLICATION NUMBER: US 60/208,802			
; PRIOR FILING DATE: 2000-06-02			
; NUMBER OF SEQ ID NOS: 78			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 75			
; LENGTH: 1497			
; TYPE: DNA			
; ORGANISM: Caenorhabditis elegans			
US-09-872-523-75			
Query Match			
99.9%; Score 1495.4; DB 10; Length 1497;			

Best Local Similarity 99.9%: Pred. No. 0; Matches 1496; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	1	ATGTCGAAATTTTCGAAAAATGTGCAGAGCTAAACAAAAATTCGACAGAAAACTCGATAG	60
Db	1	ATGTCGAAATTTTCGAAAAATGTGCAGAGCTAAACAAAAATTCGACAGAAAACTCGATAG	60
QY	61	ACCTACTTGTGGGAATCCTATTACATCAGTTCGAGAAAGAAAACTCTTTCAATCC	120
Db	61	ACCTACTTGTGGGAATCCTATTACATCAGTTCGAGAAAGAAAACTCTTTCAATCC	120
QY	121	GTTGAGACATTCATCGTAACCTACAGTTAAATTTTAAACGAATCGTAAGGAAGAGTT	180
Db	121	GTTGAGACATTCATCGTAACCTACAGTTAAATTTTAAACGAATCGTAAGGAAGAGTT	180
QY	181	ATCTTCGAACAGTGTCCATGATTATGACAAGAATCGCATTCGATTCAAGTCAGATGG	240
Db	181	ATCTTCGAACAGTGTCCATGATTATGACAAGAATCGCATTCGATTCAAGTCAGATGG	240
QY	241	TTTGACAGAAATGAAAAAGTTGGCGGATACAGAGTTCTGGCTCAGTTTATCGAGCTGAC	300
Db	241	TTTGACAGAAATGAAAAAGTTGGCGGATACAGAGTTCTGGCTCAGTTTATCGAGCTGAC	300
QY	301	ACGAAATTTTGGCTCAATATTTTATTCGAGCATATGTTGGTTGGCAACGCCGCAATG	360
Db	301	ACGAAATTTTGGCTCAATATTTTATTCGAGCATATGTTGGTTGGCAACGCCGCAATG	360
QY	361	AGTATCCCAATATGAGTAAAAATTTGTATCTCCGGCGCTTGCATCAACGAAGATAC	420
Db	361	AGTATCCCAATATGAGTAAAAATTTGTATCTCTCGCCGCTTGCATCAACGAAGATAC	420
QY	421	CAAAATGATATGTAAATTTATGTAATATATTCATGATGGCGAAATCGTCGCCAACT	480
Db	421	CAAAATGATATGTAAATTTATGTAATATATTCATGATGGCGAAATCGTCGCCAACT	480
QY	481	TGCGTGTCTCAAAATTCGATGAAGGGAAGGCTCTCTAAGCAAGCATCGTTAAAGTT	540
Db	481	TGCGTGTCTCAAAATTCGATGAAGGGAAGGCTCTCTAAGCAAGCATCGTTAAAGTT	540
QY	541	GGACAAGCTCTTGAACATTTTAAATTTATTCCTACTGAAATACGGCTAGCGCGAAT	600
Db	541	GGACAAGCTCTTGAACATTTTAAATTTATTCCTACTGAAATACGGCTAGCGCGAAT	600
QY	601	CAAGAAATATGTGACGACGAGATGATATCTATCACAAAGAAAGACTTCCGGAATCG	660
Db	601	CAAGAAATATGTGACGACGAGATGATATCTATCACAAAGAAAGACTTCCGGAATCG	660
QY	661	CTTCCAGATGAGATGACGACAGCAAGTCTTAGCTGTGATCCATATTTGGATGAC	720
Db	661	CTTCCAGATGAGATGACGACAGCAAGTCTTAGCTGTGATCCATATTTGGATGAC	720
QY	721	GAGGGAACCTCTTCATATTTCCGTGGTGATTTGCAGACGTCATGTGATCACTAAAT	780
Db	721	GAGGGAACCTCTTCATATTTCCGTGGTGATTTGCAGACGTCATGTGATCACTAAAT	780
QY	781	GCGAAGAAAGAAATATATTGAGCACACAAATTAATTTGCTCAAGCAATTAATAATGAGAA	840
Db	781	GCGAAGAAAGAAATATATTGAGCACACAAATTAATTTGCTCAAGCAATTAATAATGAGAA	840
QY	841	AATCCAGATATGACTCAGACGAGCTCACAATTTGATCAATTTAGCAAAAAGATCCAAATGAT	900
Db	841	AATCCAGATATGACTCAGACGAGCTCACAATTTGATCAATTTAGCAAAAAGATCCAAATGAT	900
QY	901	CCCATGATTTGGAGAAAAGTTAAGGTGGGAACAAAAGTTTGAGCTCATCGACCCCTTGCT	960
Db	901	CCCATGATTTGGAGAAAAGTTAAGGTGGGAACAAAAGTTTGAGCTCATCGACCCCTTGCT	960
QY	961	CAGCAATTCATTAACCTCCACGTCGCTCGATTCCTCAAAATTTTGCAAAACCTGAAGATAT	1020
Db	961	CAGCAATTCATTAACCTCCACGTCGCTCGATTCCTCAAAATTTTGCAAAACCTGAAGATAT	1020
QY	1021	CTTATGTGGGAATGATGTCAGATGACATTGAAGACAGTTTCTTATTCATATCAAT	1080
Db	1021	CTTATGTGGGAATGATGTCAGATGACATTGAAGACAGTTTCTTATTCATATCAAT	1080

Db	1021	CTTATTGTGGGAATGCAATGGATGCTCCAGATGACCTTGGAAGACACTTTTCCATTCAATCAAT	1080
Qy	1081	AATACATTATATGTTCCCGAGTTGGTTATGCGGAAAGATATAATTTGGAACCTGTGCCCA	1140
Db	1081	AATACATTATATGTTCCCGAGTTGGTTATGCGGAAAGATATAATTTGGAACCTGTGCCCA	1140
Qy	1141	GATAGATTCAAGGACATTCAGATGGGATGAATACTTTGGAGAAAGAAATCTGCAGAAACC	1200
Db	1141	GATAGATTCAAGGACATTCAGATGGGATGAATACTTTGGAGAAAGAAATCTGCAGAAACC	1200
Qy	1201	CTACCGCTTGACTGTTCACAGCCAAATGCTTCCCAAGAGAGATTAGACAAATTTAAGSTA	1260
Db	1201	CTACCGCTTGACTGTTCACAGCCAAATGCTTCCCAAGAGAGATTAGACAAATTTAAGSTA	1260
Qy	1261	ATTCTGATTTTCCAAACGGGTAGGACATCGCCTTGAAAGCTGCTGACATGTGGAATAATCAG	1320
Db	1261	ATTCTGATTTTCCAAACGGGTAGGACATCGCCTTGAAAGCTGCTGACATGTGGAATAATCAG	1320
Qy	1321	TTTTTTTGTCACAGTTACAGTGAATAATCACTTCATGGAAGACCTGATTAATGTCAATTTTGAC	1380
Db	1321	TTTTTTTGTCACAGTTACAGTGAATAATCACTTCATGGAAGACCTGATTAATGTCAATTTTGAC	1380
Qy	1381	GGCTGGGATGAAGAAATTTGATGAACCTGATGATGTGCACTCCATGATATTTCAACGATA	1440
Db	1381	GGCTGGGATGAAGAAATTTGATGAACCTGATGATGTGCACTCCATGATATTTCAACGATA	1440
Qy	1441	GGATGGTGTGAAGCGCACAGTTATGTTCTACACCTCCGAAAAAGTACAACATATATGA	1497
Db	1441	GGATGGTGTGAAGCGCACAGTTATGTTCTACACCTCCGAAAAAGTACAACATATATGA	1497

```

RESULT 5
US-09-872-523-76
; Sequence 76, Application US/09872523
; Patent No. US20020137906A1
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Davison, Ewa M.
; APPLICANT: Lu, Xiaowei
; TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
; FILE REFERENCE: 0197/53602
; CURRENT APPLICATION NUMBER: US/09/872,523
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/208,802
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 2307
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-872-523-76

```

Query Match	73.7%	Score 1103;	DB 10;	Length 2307;
Best Local Similarity	81.3%	Pred. No. 1.3e-284;		
Matches 1497; Conservative	0;	Mismatches 0;	Indels 344;	Gaps 5;

QY	1	ATGCTGGAATTTGGAATAATGTGAGCTAACAAAAATCGGACAGAAAACCTGATPAG	60
Db	467	ATGCTGGAATTTGGAATAATGTGAGCTAACAAAAATCGGACAGAAAACCTGATPAG	526
QY	61	ACCTACTTGTGGGAATCCTATTTCATCAGCTCGAGAAAGGAAACCTCTTCATTCCA	120
Db	527	ACCTACTTGTGGGAATCCTATTTCATCAGCTCGAGAAAGGAAACCTCTTCATTCCA	586
QY	121	GTTCGAGCAATTCATCGTAACCTTACAGTTAATTTTAAACGAAATCGTGAAGGAAGATT	180
Db	587	GTTCGAGCAATTCATCGTAACCTTACAGTTAATTTTAAACGAAATCGTGAAGGAAGATT	646
QY	181	ATC-----TTGCAACA	192
	111		
	647	ATCTGTGATCATATTGTTCGTAATACGGTTTAAATACAAATTTTGATGTGGAACA	706

OY	193	GTGTCATGATTAATGACAGACACTCGGATTCGATTCAGTCAGATGGTTTGACGAATT	252
Db	707	GTGTCATGATTAATGACAGAACACTCGCATTCGATTCAGTCAGATGGTTTGACGAATT	766
OY	253	GAAAAGCTTCGGATACAGAGTCTGGCTCAGATTATACGAGCTGCACGAATTTTGG	312
Db	767	GAAAAGCTTCGGATACAGAGTCTGGCTCAGATTATACGAGCTGCACGAATTTTGG	826
OY	313	CTCATATTTTATCGACGATGTTGGTTGGCAAA-----	350
Db	827	CTCATATTTTATCGACGATATGTTTGGTTGGCAAAAGTAGTTTGGACGCTCAGCTCTT	886
OY	351	-----	350
Db	887	TCTACTATTTCAATAATAATAAGTGTCTGTTACATAAATTTAGAGACAATCGTATTA	946
OY	351	-----CGCGCGCAATGAGAT	366
Db	947	AAACTTGAACATTTGTATATATGTAATAATTTGACACTTTACAGCGCGCCAGATGATGAT	1006
OY	367	CCCAATATGATAAAATTTGATATGCTCCGCGCTTGCAATCAAGAGAATTACCAAAAT	426
Db	1007	CCCAATATGATAAAATTTGATATGCTCCGCGCTTGCAATCAAGAGAATTACCAAAAT	1066
OY	427	GATATGTAATTTATGTAAT-----	447
Db	1067	GATATGTAATTTATGTAATGTAAGTTGTTTTTTTCCGAATTAATGTAATATCATCT	1126
OY	448	-----AATTGCATTTGATGGCAAAATCGTCCGCCCAACTTGGCTGTCTCCAAAT	496

112	TCACGATTCACAAATTCGATTCATGCGGAAATTCGTCGCGCAATTCGCTGCTCCAAATTC	110
QY	497 TCGATGAAAGGAAGGCGCTCCCTCCAAAGCAAGCATGCTTTCAAGTTGGACAAACGTCTTGAC	556
Db	1187 TCGATGAAAGGAAGGCGCTCCCTCCAAAGCAAGCATGCTTTCAAGTTGGACAAACGTCTTGAC	1246
QY	557 TATTAATAATTTATCCAAATTCCTACTGTAATAGCGCGGAATTCGAATTAATGTGGAC	616
Db	1247 TATTAATAATTTATCCAAATTCCTACTGTAATAGCGCGGAATTCGAATTAATGTGGAC	1306
QY	617 GAGCAATGAATGTATCTATTCACAAAGAAAGACTTTCCCAATGCGCTTCAGATGCGATG	676
Db	1307 GAGCAATGAATGTATCTATTCACAAAGAAAGACTTTCCCAATGCGCTTCAGATGCGATG	1366
QY	677 ACGACAGACAGCTCTTATAGTCTGTGGATGTCGAATTTGGATATGACGAGGGAAGCTTCTCA	736
Db	1367 ACGACAGACAGCTCTTATAGTCTGTGGATGTCGAATTTGGATATGACGAGGGAAGCTTCTCA	1426
QY	737 TATTTCCCTGTTGGATTTTGCAGCAGCTCAATGATATCACTAATTCGCAAAAAGGAATTA	796
Db	1427 TATTTCCCTGTTGGATTTTGCAGCAGCTCAATGATATCACTAATTCGCAAAAAGGAATTA	1486
QY	797 TTGGAGCACACAAATTAATTTGCTCAGCAATTAATAAATGGAGAAATTCAGATATATGCT	856
Db	1487 TTGGAGCACACAAATTAATTTGCTCAGCAATTAATAAATGGAGAAATTCAGATATATGCT	1546
QY	857 CAGACGACGTCACATTTGATCAATTAATAGCAAAAGATCCAAATTCATCCATGATTTGGAGAA	916
Db	1547 CAGACGACGTCACATTTGATCAATTAATAGCAAAAGATCCAAATTCATCCATGATTTGGAGAA	1606
QY	917 AAGTTAAGGTTGGACAAAGTTTGAAGTCATCGACCCCTTGGGTCAGCAATTCAAATAC	976
Db	1607 AAGTTAAGGTTGGACAAAGTTTGAAGTCATCGACCCCTTGGGTCAGCAATTCAAATAC	1666
QY	977 TCCACGTCGCTTCATATTCCTCAAAATTTTGGCAAAACGTGAAGATATCTTATTTGGGAATGG	1036
Db	1667 TCCACGTCGCTTCATATTCCTCAAAATTTTGGCAAAACGTGAAGATATCTTATTTGGGAATGG	1726
QY	1037 ATGATCCAGATGCACTTGAAGACAGTTTCTATATTCATATCAATTAATATACATTTATGTTCC	1096
Db	1727 ATGATCCAGATGCACTTGAAGACAGTTTCTATATTCATATCAATTAATATACATTTATGTTCC	1786
QY	1097 CAGTTGGTATAGCGGAAAGATATATTTGGAAATCTGTTCCGCGAGATGAGTTCAAAGGAA	1156

Db	1787	CAGTTGGTTATCGGAAAAGTATATTTGGAACTTTGTCGCCGAGATGATTCAAAGGAA	1846
Qy	1157	CATTCCAGATGGAGAACTACTTGGAGAAAGAACTGCGAAGAACCCCTACCGTTGACTTTGT	1216
Db	1847	CATTCCAGATGGAGAACTACTTGGAGAAAGAACTGCGAAGAACCCCTACCGTTGACTTTGT	1906
Qy	1217	TCAAGCCAAATGCCCTTCCCAAGAGAGATTAGACAAATTTTAAGTAAATTCGATTTTCCAAAC	1278
Db	1907	TCAAGCCAAATGCCCTTCCCAAGAGAGATTAGACAAATTTTAAGTAAATTCGATTTTCCAAAC	1966
Qy	1277	G-----	1277
Db	1967	GGGTTTGTATTAATCGTTTGAGATTTGTTTCACTATTAATTAAGTTATTCATAATGTTTCTT	2028
Qy	1278	-----GGTAGAGACTACGCCCTTGAAGCTGCTGACATGTGTGAAAATCAGTTTAATTTGTC	1330
Db	2027	GTTTAAAGTAGAGACTACGCCCTTGAAGCTGCTGACATGTGTGAAAATCAGTTTAATTTGTC	2086
Qy	1331	CAGCTACAGTGAATTCAGCTTCATGGAAGACATGATTAATGTCAATTTTGCAGCGCTGGAGATG	1390
Db	2087	CAGCTACAGTGAATTCAGCTTCATGGAAGACATGATTAATGTCAATTTTGCAGCGCTGGAGATG	2148
Qy	1391	AAGAAATTTGATGAACCTGTATATGTGGA-----	1418
Db	2147	AAGAAATTTGATGAACCTGTATATGTGGAAGTTTATCATGACCGAAGACATTTTTC	2206
Qy	1419	-----CTCCCATGATTAATTCACCATGATGAGATGATGTTGAAGCGC	1456
Db	2207	AATGAAAATTCATATCATTTTCAGCTCCCATGATATTTCTACCATATAGATGATGTTGAAGCGC	2266
Qy	1457	ACAGTTATGTTCTACAACTCCGAAAAAGTACAACTATTGA	1497
Db	2267	ACAGTTATGTTCTACAACTCCGAAAAAGTACAACTATTGA	2307
RESULT 6			
US-09-872-523-77			
Sequence 77, Application US/09872523			
Patent No. US20020137906A1			
GENERAL INFORMATION:			
APPLICANT: Horvitz, H. Robert			
APPLICANT: Davison, Ewa M.			
APPLICANT: Lu, Xiaowei			
TITLE OF INVENTION: A Tumor Suppressor Pathway In C. Elegans			
FILE REFERENCE: 01997/536002			
CURRENT APPLICATION NUMBER: US/09/872,523			
CURRENT FILING DATE: 2001-06-01			
PRIOR APPLICATION NUMBER: US 60/208,802			
PRIOR FILING DATE: 2000-06-02			
NUMBER OF SEQ ID NOS: 78			
SOFTWARE: FastSeq for Windows Version 4.0			
SEQ ID NO 77			
LENGTH: 2307			
TYPE: DNA			
ORGANISM: Caenorhabditis elegans			
US-09-872-523-77			
Query Match			
Best Local Similarity 73.7%; Score 1103; DB 10; Length 2307;			
Matches 1497; Conservativity 81.3%; Pred. No. 1.3e-282;			
Matches 1497; Conservativity 0; Mismatches 0; Indels 344; Gaps			
Qy	1	ATGCTGAAATTTCTGAAAATTTGTCAGACTTAACAACAAAATATCGACAGAAAACCTCTTCAATTC	60
Db	467	ATGCTGAAATTTCTGAAAATTTGTCAGACTTAACAACAAAATATCGACAGAAAACCTCTTCAATTC	526
Qy	61	ACCTACTGTGGGAATCCCTATTTACATCAGTTCGAGAAAGGAAAAAATCTTCTTCAATTC	120
Db	527	ACCTACTGTGGGAATCCCTATTTACATCAGTTCGAGAAAGGAAAAAATCTTCTTCAATTC	586
Qy	121	GTTGAGACATTCATCGTACCTTACAGTTAATTTTAAACGAATGCGTGAAGAGAGATT	180
Db	587	GTTGAGACATTCATCGTACCTTACAGTTAATTTTAAACGAATGCGTGAAGAGAGAGATT	646

[illegible]

Qy	1037	ATGTCAGATGCACTTGAGACGCTTTCATATCAATCAAAATACATTAATGCTCC	1096
Db	1727	ATGGTCCAGATGCACTTGAGACGCTTTCATATCAATCAAAATACATTAATGCTCC	1786
Qy	1097	CAGTTGGTTATGCGGAAAGATATATTTGGAACTGTTCCGCGAGATGATTCAAAGAA	1156
Db	1787	CAGTTGGTTATGCGGAAAGATATATTTGGAACTGTTCCGCGAGATGATTCAAAGAA	1846
Qy	1157	CATTACATGGGATGAATTACTTGAGAAAGAACTGCAGAAACCTTACCGCTGACTTGT	1216
Db	1847	CATTACATGGGATGAATTACTTGAGAAAGAACTGCAGAAACCTTACCGCTGACTTGT	1906
Qy	1217	TCAAGCCCAATGCCCTCCCAAGAGAGATTAGCAAAATTAAGGTAACTGTGATTTCCAAAC	1276
Db	1907	TCAAGCCCAATGCCCTCCCAAGAGAGATTAGCAAAATTAAGGTAACTGTGATTTCCAAAC	1966
Qy	1277	G-----	1277
Db	1967	GGGTGTTTATCGTTTGAGATTGTTCCACTATTAATAGTAATCATATATGTTTCTT	2026
Qy	1278	-----GTTAGAGCTACGCTTGAAGCTGTGACATGTGTGAAATACAGTTTATTTGTC	1330
Db	2027	GTTTAAAGTAGAGACTACGCTTGAAGCTGTGACATGTGTGAAATACAGTTTATTTGTC	2086
Qy	1331	CAGCTACAGTGAATCAATCAGTTTCATGGAAGACATGATTAATGTCAATTTGCAGCGCTGGGATG	1390
Db	2087	CAGCTACAGTGAATCAATCAGTTTCATGGAAGACATGATTAATGTCAATTTGCAGCGCTGGGATG	2146
Qy	1391	AAGAAATTTGATGAACCTGTATGATGTG-----	1417
Db	2147	AAGAAATTTGATGAACCTGTATGATGTGAGTTTATCATGACCGAAGACATTTTTC	2206
Qy	1418	-----ACGCCATGATATTTACAGATAGATGTGTGAACGCC	1456
Db	2207	AATGAAATTTCTATCATTTCAACGTCACCATGATATTTTACGATAGGATGTGTGAACGCC	2266
Qy	1457	ACAGTTATGTTCTACAACTCCGAAAAAGTACAACTATTGA	1497
Db	2267	ACAGTTATGTTCTCTCAACACTCCGAAAAAGTACAACTATTGA	2307
RESULT 7			
US-09-872-523-78			
: Sequence 78, Application US/09872523			
: Patent No. US20020137906A1			
: GENERAL INFORMATION:			
: APPLICANT: HORVITZ, H. ROBERT			
: APPLICANT: DAVIDSON, EWA M.			
: TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans			
: FILE REFERENCE: 01997/536002			
: CURRENT APPLICATION NUMBER: US/09/872,523			
: PRIOR FILING DATE: 2001-06-01			
: PRIOR APPLICATION NUMBER: US 60/208,802			
: NUMBER OF SEQ ID NOS: 78			
: SOFTWARE: FastSeq for Windows Version 4.0			
: SEQ ID NO 78			
: LENGTH: 2307			
: TYPE: DNA			
: ORGANISM: Caenorhabditis elegans			
: US-09-872-523-78			

Query Match	73.7%	Score 1103;	DB 10;	Length 2307;
Best Local Similarity	81.3%	Pred. No. 1.3e-282;		
Matches 1497;	Conservative	0;	Mismatches	0;
			Indels	344;
			Gaps	5;
QY	1	ATGCTCAAAATTTTCGAAATTTGTCAGAGCTACAAAAATTCGACAGAAAACTCGATAAG	60	
Db	467	ATGCTCAAAATTTTCGAAATTTGTCAGAGCTACAAAAATTCGACAGAAAACTCGATAAG	526	
QY	61	ACTACTTGTGGGATCCCAATTACATCAGTTCGAGAAAGAAAACTCTTCATTTCA	120	

Db	527	ACCTACTTGTGGGAATCCATTTTACTTACGTACGTTCCGAAAAAGAAAAAATCTTTCATTC	586
Qy	121	GTTCGAAGCATTCATCCTGTAACCTTACAGTTAAATTTTAAAGCATGCGTGAAGAGAGATT	180
Db	587	GTTCGAAGCATTCATCCTGTAACCTTACAGTTAAATTTTAAAGCATGCGTGAAGAGAGATT	646
Qy	181	ATC-----TTCGAAACA	192
Db	647	ATCGAGAGTTCATTTGTTGCTAAATCGGTTTTTAAATACAAATTTTGTAGTTCGAAACA	706
Qy	193	GTGTCATGATTTATGACAAAGAACTGCAATTCGATTCGAAGTCAGATGTTTGCACGAAT	252
Db	707	GTGTCATGATTTATGACAAAGAACTGCAATTCGATTCGAAGTCAGATGTTTGCACGAAT	766
Qy	253	GAAGAAAGTTTGGCGATACAGAGTTCTGGCTCAGTTTATCGAGCTGACACGAAATTTTGG	312
Db	767	GAAGAAAGTTTGGCGATACAGAGTTCTGGCTCAGTTTATCGAGCTGACACGAAATTTTGG	826
Qy	313	CTCATATTTTATCGAGCATATGTTTGGTTGGCAA-----	350
Db	827	CTCATATTTTATCGAGCATATGTTTGGTTGGCAA-----	886
Qy	351	-----	350
Db	887	TCTACTATTCTAAATTAATTAATGTTCTGTCTACATTAATTTCTAGAGACAAATCGTATTA	946
Qy	351	-----CGCGCAATGAGTGAT	366
Db	947	AAACTTCGAACATTTGTATTAATAGTAAATTTGAACATTTTCAGCGCGCAATGAGTGAT	1006
Qy	367	CCCATATGATGATTAATTTGTATATGCTCCGCGCTTCGAATCAACGAAGAAATACCAAAAT	426
Db	1007	CCCATATGATGATTAATTTGTATATGCTCCGCGCTTCGAATCAACGAAGAAATACCAAAAT	1066
Qy	427	GATATGTTAAATTTATGTAAT-----	447
Db	1067	GATATGTTAAATTTATGTAATTTGTAATGTAAGTTGTTTTTCCGAATTTATGTTAAATATCATCT	1126
Qy	448	-----AATTCGATTTGATGGCGAAATCGTGGCGCAAACTTGGCTCTCCAAAT	496
Db	1127	CACAACCTTCAAAATTCATTTGATGGCGAAATCGTGGCGCAAACTTGGCTCTCCAAAT	1186
Qy	497	TGATGTAAGGGAAGGCTCTCTCTAAGCAAGCATGTTTCAAAAGTTGGACAACGTTTAAAC	556
Db	1187	TGATGTAAGGGAAGGCTCTCTCTAAGCAAGCATGTTTCAAAAGTTGGACAACGTTTAAAC	1246
Qy	557	TATTTAAATTTTCCATTTCTACTGGAATACGGCTAGCGGAATTCAGAAATATGTTGAC	616
Db	1247	TATTTAAATTTTCCATTTCTACTGGAATACGGCTAGCGGAATTCAGAAATATGTTGAC	1306
Qy	617	GACGATGATGATATCTATCACAAAGAAAGACTTCCGGAATGCGTCCAGATGCAGATG	676
Db	1307	GACGATGATGATATCTATCACAAAGAAAGACTTCCGGAATGCGTCCAGATGCAGATG	1366
Qy	677	ACGACAGACAGCTCTTACGCTTGATCTCAATATTGGATAGACAGGGAAGCTTCTTCA	736
Db	1367	ACGACAGACAGCTCTTACGCTTGATCTCAATATTGGATAGACAGGGAAGCTTCTTCA	1426
Qy	737	TATTTCCCTGTTGGATTTGACAGCAAGTCAATGATATCACTAAATTTGGAAAAAGGAATATA	796
Db	1427	TATTTCCCTGTTGGATTTGACAGCAAGTCAATGATATCACTAAATTTGGAAAAAGGAATATA	1486
Qy	797	TTCAGACACAAATTAATTTGCTCAAGCAATAAAAAATGAGAAAAATCCAAAGATATGACT	856
Db	1487	TTCAGACACCAATTAATTTGCTCAAGCAATAAAAAATGAGAAAAATCCAAAGATATGACT	1546
Qy	857	CAGACAGCTCACATTTGATCAATTAAGCAAAAGATCCAATTGATCCCATGATTTGGAAA	916
Db	1547	CAGACAGCTCACATTTGATCAATTAAGCAAAAGATCCAATTGATCCCATGATTTGGAAA	1606
Qy	917	AAGTTAAGTTGGACAAAAGTTTGGAGCTATGAGACCCCTTGGCTCAGCAATTAATTAAC	976
Db	1607	AAGTTAAGTTGGACAAAAGTTTGGAGCTATGAGACCCCTTGGCTCAGCAATTAATTAAC	1666

```
OY 977 TCACGCTGCTTCATTCCTCAAAATTTTGCAAACTGAGAGATATCTATTGTGGGAATGG 1036
|||||
1667 TCACGCTGCTTCATTCCTCAAAATTTTGCAAACTGAGAGATATCTATTGTGGGAATGG 1726
OY 1037 ATGCTCCAGATGCACTGGAAGACGTTTCTATTATTCATATCAATATATCATTTATGTTCC 1096
|||||
1727 ATGCTCCAGATGCACTGGAAGACGTTTCTATTATTCATATCAATATATCATTTATGTTCC 1786
OY 1097 CAGTTGGTTATGCGGAAAAGATATATTTGGAACCTGTTCCCGCAGATGAGTTCAAGGAA 1156
|||||
1787 CAGTTGGTTATGCGGAAAAGATATATTTGGAACCTGTTCCCGCAGATGAGTTCAAGGAA 1846
OY 1157 CATTCAAGTGGAGTAATCTTGTGAGAAAAGATCTGAGAAAACCTTACCCTTACCTTGT 1216
|||||
1847 CATTCAAGTGGAGTAATCTTGTGAGAAAAGATCTGAGAAAACCTTACCCTTACCTTGT 1906
OY 1217 TCACGCCAATGCTTCCCAAGAGATTTAGACAAATTAAGCTATCTGATTTCCAAAC 1276
|||||
1907 TCACGCCAATGCTTCCCAAGAGATTTAGACAAATTAAGCTATCTGATTTCCAAAC 1966
OY 1277 G----- 1277
1967 GGGTTGTTTATTCGTTTGATTTGTTTCTACTATTAATAGTTATTCATATTTGTTCTT 2026
OY 1278 -----GTAAGACTACGCTTGAAGCTGCTGACATGTGGAATTCAGTTTATTTGTC 1330
|||||
2027 GTTTTAAGGTAGGACTACGCTTGAAGCTGCTGACATGTGGAATTCAGTTTATTTGTC 2086
OY 1331 CAGCTACAGTAATTCAGTTGATGGAAGACTGATTAATGTAATTTGACGCGGCGGAGTG 1390
|||||
2087 CAGCTACAGTAATTCAGTTGATGGAAGACTGATTAATGTAATTTGACGCGGCGGAGTG 2146
OY 1391 AAGAATTTGATGAACCTGTATGATGTGA----- 1418
|||||
2147 AAGAATTTGATGAACCTGTATGATGTGAAGTTTATCATGACCGAAGACATTTTCTC 2206
OY 1419 -----CTCCATGATATTTCTACCGATGAGATGCGTGAAGCGC 1456
|||||
2207 AATGAAAATTCATTCATTCAGCTCCATGATATTCATCCATAGGATGAGTGTGAAGCGC 2266
OY 1457 ACAGTTATGTTCTACAACTCCGAAAAAGTACAACTATTGA 1497
|||||
2267 ACAGTTATGTTCTACAACTCCGAAAAAGTACAACTATTGA 2307
Db
```

RESULT 8

```
US-09-801-574-11
; Sequence 11, Application US/09801574
; Patent No. US20020081592A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Beijing Jeremy
; TITLE OF INVENTION: Reproduction-Specific Genes
; FILE REFERENCE: 0399,2007-002
; CURRENT APPLICATION NUMBER: US-09/801,574
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: 60/187,518
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: 60/261,557
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 2969
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-801-574-11
```

```
Query Match 3.9%; Score 58.8; DB 9; Length 2969;
Best Local Similarity 55.3%; Pred. No. 6.6e-05;
Matches 114; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
```

```
OY 1279 GTAGACTACGCCCTTGAACTGCTGACATGTGTAATAATAGTTATTATGTCACGCTACA 1338
|||||
801 GTTGGATGTAAGATGGAAGCTGTCGATAGAAAAATCCATTATGATCTGTCGACACA 860
OY 1339 GTGAATCAGTTCATGGAAGACGTAATATGTCAATTTTCGACGCGTGGATGAAGAAATTT 1398
|||||
861 ATTGAGCTGTCTGTGAGATCACTCAATCTATCTCTTTGATGATGATGAGTGAAGCATTT 920
OY 1399 GATGACCTGATGATGATGAGATCCCATGATTAATTCATCCGATAGAGATGATGTAAGCGCAC 1458
|||||
921 GATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 980
OY 1459 AGTTATGTTCTACAACTCCGAAAAA 1484
|||||
981 GGAGATGATTAATACAGCCACGAGAAA 1006
Db
```

RESULT 9

```
US-09-789-919-55
; Sequence 55, Application US/09789919
; Patent No. US2002006485A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Moore, Kateri
; TITLE OF INVENTION: GENES THAT REGULATE HEMATOPOIETIC BLOOD FORMING STEM
; FILE REFERENCE: 2275-1-005
; CURRENT APPLICATION NUMBER: US/09/789,919
; CURRENT FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 55
; LENGTH: 3590
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-789-919-55
```

```
Query Match 3.8%; Score 57.4; DB 9; Length 3590;
Best Local Similarity 55.2%; Pred. No. 0.00017;
Matches 112; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
```

```
OY 1276 CGGCTAGGACTACGCCCTTGAACTGCTGACATGTGTAATAATGATTTATTTGTCACGCT 1335
|||||
1745 CGGCTAGGAAATGAATTAAGAGCTGTAGATCTATGAGCGCAGGTTAATATGTGTAGGCC 1804
OY 1336 ACAGTGAATCAGTTTCATGGAAGACGTAATATGTCATTTTCGACGCGTGGATGAAGAA 1395
|||||
1805 ACAGTGAATCAGTTTCATGGAAGACGTAATATGTCATTTTCGACGCGTGGATGAAGAA 1864
OY 1396 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1455
|||||
1865 TATGACCACTGCTGAGAGCTGTGAGTCTCCCTGACCTCTATCTGTAGGGTGTGACGTTA 1924
OY 1456 CACAGTTATGTTCTACAACTCC 1478
|||||
1925 ACTGATATCAACTACAGCCTCC 1947
Db
```

RESULT 10

```
US-09-764-868-396
; Sequence 396, Application US/09764868
; Patent No. US2002016871A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 396
; LENGTH: 2403
```

```

: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (1)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (10)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (18)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (41)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (2398)
: OTHER INFORMATION: n equals a,t,g, or c
US-09-764-868-396
```

```

Query Match          3.7%; Score 55.2; DB 10; Length 2403;
Best Local Similarity 55.9%; Pred. No. 0.00053;
Matches 105; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
```

```

QY 1291 CTTGAAGCTGCTGACATGTGTGAAATCAGTTATTGTCACAGTACAGTAATCAGTT 1350
      ||||| || ||||| | | ||||| | ||||| | ||||| |
DB 700 CTTGAGAGTGTAGACAAAGAACCCCTATGTTATTAGAGTACAGTGTGGACAGACA 759
      ||||| || ||||| | | ||||| | ||||| | ||||| |
QY 1351 CATGAAGACTGATAAATGTCATTTTCGACGGCTGGATGAGAAATTTGATGAACTGTAT 1410
      ||||| || ||||| | | ||||| | ||||| | ||||| |
DB 760 GATGATCACCGGGGTAAGTTCCTTCTGATGCTGGACAAATGCTATGATTCCTGATA 819
      ||||| || ||||| | | ||||| | ||||| | ||||| |
QY 1411 GATGTGACCTCCATGATTTCTACGATAGATGATGATGATGATGATGATGATGATGAT 1470
      ||||| || ||||| | | ||||| | ||||| | ||||| |
DB 820 GATGACGATTTCTCTGATATTCACCTGTAGGCTGTGTTCAAAAACAGACATCCCTT 879
      ||||| || ||||| | | ||||| | ||||| | ||||| |
QY 1471 CAACCTCC 1478
      ||||| || ||||| | | ||||| | ||||| | ||||| |
DB 880 CAGCTCC 887
```

```

RESULT 11
US-10-204-887-14
: Sequence 14, Application US/10204887
: Publication No. US20030124569A1
: GENERAL INFORMATION:
: APPLICANT: INCYTE GENOMICS, INC.
: APPLICANT: PANZER, Scott R.
: APPLICANT: SPIRO, Peter A.
: APPLICANT: BANYVILLE, Steven C.
: APPLICANT: SHAH, Purvi
: APPLICANT: CHAIUP, Michael S.
: APPLICANT: CHANG, Simon C.
: APPLICANT: CHEN, Alice
: APPLICANT: D/SA, Steven A.
: APPLICANT: AMSHRY, Stefan
: APPLICANT: DAHL, Christopher R.
: APPLICANT: DAM, Tam C.
: APPLICANT: DANIELS, Susan E.
: APPLICANT: DUFOR, Gerard E.
: APPLICANT: FLORES, Vincent
: APPLICANT: FONG, Willy T.
: APPLICANT: GREENAWALT, Lila B.
: APPLICANT: HILLMAN, Jennifer L.
: APPLICANT: JONES, Anissa L.
: APPLICANT: LIU, Tommy F.
: APPLICANT: ROSEBERRY, Ann M.
: APPLICANT: ROSEN, Bruce H.
: APPLICANT: RUSSO, Frank D.
: APPLICANT: STOCKDREHER, Theresa K.
: APPLICANT: DAFEO, Abel
: APPLICANT: WRIGHT, Rachel J.
: APPLICANT: YAP, Pierre E.
```

```

: APPLICANT: YU, Jimmy Y.
: APPLICANT: BRADLEY, Diana L.
: APPLICANT: BRATCHER, Shawn R.
: APPLICANT: CHEN, Wensheng
: APPLICANT: COHEN, Howard J.
: APPLICANT: HODGSON, David M.
: APPLICANT: LINCOLN, Stephen E.
: TITLE OF INVENTION: SECRETORY MOLECULES
: FILE REFERENCE: PT-1134 PCT
: CURRENT APPLICATION NUMBER: US/10/204,887
: CURRENT FILING DATE: 2002-08-21
: PRIOR APPLICATION NUMBER: 60/185,215; 60/185,216; 60/205,232; 60/205,233; 60/205,2
  60/205,324; 60/205,286
: PRIOR FILING DATE: 2000-02-24; 2000-02-24; 2000-05-16; 2000-05-17; 2000-05-17;
  2000-05-17; 2000-05-17
: NUMBER OF SEQ ID NOS: 159
: SOFTWARE: PERL Program
: SEQ ID NO 14
: LENGTH: 1035
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: OTHER INFORMATION: Incyte ID No. US20030124569A1 LI:228319.1:2000PEB01
US-10-204-887-14
```

```

Query Match          3.6%; Score 53.4; DB 14; Length 1035;
Best Local Similarity 55.0%; Pred. No. 0.00099;
Matches 105; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
```

```

QY 1291 CTTGAAGCTGCTGACATGTGTGAAATCAGTTATTGTCACAGTACAGTAATCAGTT 1350
      ||||| || ||||| | | ||||| | ||||| | ||||| |
DB 815 CTTGAAGTGTGATTAACGAAACCCAGTTATTCGTTCTACGATTTGATGATTT 874
      ||||| || ||||| | | ||||| | ||||| | ||||| |
QY 1351 CATGAAGACTGATAAATGTCATTTTCGACGGCTGGATGAGAAATTTGATGAACTGTAT 1410
      ||||| || ||||| | | ||||| | ||||| | ||||| |
DB 875 GATGACCAAAAGATTAAGTTTCATTTGATGTTGGACCATTAATGATGACTACTGGCTG 934
      ||||| || ||||| | | ||||| | ||||| | ||||| |
QY 1411 GATGTGACCTCCATGATTTCTACGATAGATGATGATGATGATGATGATGATGATGAT 1470
      ||||| || ||||| | | ||||| | ||||| | ||||| |
DB 935 GAGGACAGACGCCCTGATATTCACCCGATGCGATTTGTGATGTCACAGGACATCCACTG 994
      ||||| || ||||| | | ||||| | ||||| | ||||| |
QY 1471 CAACCTCCGAA 1481
      ||||| || ||||| | | ||||| | ||||| | ||||| |
DB 995 GAAGTCCCAA 1005
```

```

RESULT 12
US-09-864-761-20410
: Sequence 20410, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
  R FILE REFERENCE: Aeonica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
```

```

1  PRIOR APPLICATION NUMBER: PCT/US01/00667
2  PRIOR FILING DATE: 2001-01-30
3  PRIOR APPLICATION NUMBER: PCT/US01/00664
4  PRIOR FILING DATE: 2001-01-30
5  PRIOR APPLICATION NUMBER: PCT/US01/00669
6  PRIOR FILING DATE: 2001-01-30
7  PRIOR APPLICATION NUMBER: PCT/US01/00665
8  PRIOR FILING DATE: 2001-01-30
9  PRIOR APPLICATION NUMBER: PCT/US01/00668
10 PRIOR FILING DATE: 2001-01-30
11 PRIOR APPLICATION NUMBER: PCT/US01/00663
12 PRIOR FILING DATE: 2001-01-30
13 PRIOR APPLICATION NUMBER: PCT/US01/00662
14 PRIOR FILING DATE: 2001-01-30
15 PRIOR APPLICATION NUMBER: PCT/US01/00661
16 PRIOR FILING DATE: 2001-01-30
17 PRIOR APPLICATION NUMBER: PCT/US01/00670
18 PRIOR FILING DATE: 2001-01-30
19 PRIOR APPLICATION NUMBER: US 60/234,687
20 PRIOR FILING DATE: 2000-09-21
21 PRIOR APPLICATION NUMBER: US 09/608,408
22 PRIOR FILING DATE: 2000-06-30
23 PRIOR APPLICATION NUMBER: US 09/774,203
24 PRIOR FILING DATE: 2001-01-29
25 NUMBER OF SEQ. ID NOS: 491117
26 SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
27 SEQ ID NO 20410
28 LENGTH: 204
29 TYPE: DNA
30 ORGANISM: Homo sapiens
31 FEATURE:
32 OTHER INFORMATION: MAP TO AL110502.1
33 OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 0.93
34 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
35 OTHER INFORMATION: EXPRESSED IN B4474, SIGNAL = 1.1
36 OTHER INFORMATION: SWISSPROT HIT: Q01441, EVALU2.2,20e+00
37 OTHER INFORMATION: EST HUMAN HIT: BE256221.1, EVALU1.1,00e-111
38 OTHER INFORMATION: NT HIT: g16512641, EVALU2.1,00e-111
39
40 US-09-864 -761-20410

```


THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2003, 14:53:19 ; Search time 3439 Seconds
(without alignments)
10579.762 Million cell updates/sec

Title: US-09-872-523-6
Perfect score: 1497
Sequence: 1 atgtctgaattcttgaanaat.....cgaaaagracacattga 1497

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estinu:*
5: em_estov:*
6: em_estopl:*
7: em_estro:*
8: em_estl:*
9: gb_estl:*
10: gb_estc:*
11: gb_hic:*
12: gb_estc3:*
13: gb_estl4:*
14: gb_estl5:*
15: em_estfun:*
16: em_estrom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pin:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gssl:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	630.4	42.1	9 AU204939	AU204939 AU204939
2	599	40.0	12 BU143186	BU143186 BU143186
3	591	39.5	12 BU129361	BU129361 BU129361
4	563.4	37.6	14 CB404801	CB404801 OSTR028D3

5	559	37.3	581	12	BU101701	BU101701
6	521	34.8	557	12 <th>BU124772</th> <th>BU124772</th>	BU124772	BU124772
7	519	34.7	677	9 <th>AU216908</th> <th>AU216908</th>	AU216908	AU216908
8	510	34.1	686	12 <th>BU153087</th> <th>BU153087</th>	BU153087	BU153087
9	502.4	33.6	519	12 <th>BU115564</th> <th>BU115564</th>	BU115564	BU115564
10	366	24.4	380	9 <th>AV197399</th> <th>AV197399</th>	AV197399	AV197399
11	360	24.0	360	14 <th>CA42305</th> <th>CA42305</th>	CA42305	CA42305
12	308	20.6	340	9 <th>AU111588</th> <th>AU111588</th>	AU111588	AU111588
13	220	14.7	300	9 <th>AU116296</th> <th>AU116296</th>	AU116296	AU116296
14	215	14.4	300	9 <th>AU115428</th> <th>AU115428</th>	AU115428	AU115428
15	186.2	12.4	337	9 <th>AU112450</th> <th>AU112450</th>	AU112450	AU112450
16	80.2	5.4	377	14 <th>CA3257</th> <th>CA3257</th>	CA3257	CA3257
17	66	4.4	505	14 <th>CB385690</th> <th>CB385690</th>	CB385690	CB385690
18	65	4.3	360	14 <th>C70202</th> <th>C70202</th>	C70202	C70202
19	60.6	4.0	554	10 <th>BE684409</th> <th>BE684409</th>	BE684409	BE684409
20	59.8	4.0	291	14 <th>CB276655</th> <th>CB276655</th>	CB276655	CB276655
21	59	3.9	350	9 <th>AA874037</th> <th>AA874037</th>	AA874037	AA874037
22	59	3.9	521	13 <th>BQ552847</th> <th>BQ552847</th>	BQ552847	BQ552847
23	59	3.9	548	9 <th>AU045245</th> <th>AU045245</th>	AU045245	AU045245
24	59	3.9	605	28 <th>AZ952671</th> <th>AZ952671</th>	AZ952671	AZ952671
25	59	3.9	655	28 <th>BB068384</th> <th>BB068384</th>	BB068384	BB068384
26	59	3.9	664	13 <th>BO552848</th> <th>BO552848</th>	BO552848	BO552848
27	59	3.9	4114	11 <th>AK028503</th> <th>AK028503</th>	AK028503	AK028503
28	59	3.9	4852	11 <th>AK033238</th> <th>AK033238</th>	AK033238	AK033238
29	58.8	3.9	2938	11 <th>AK016533</th> <th>AK016533</th>	AK016533	AK016533
30	58.8	3.9	3442	11 <th>AK040521</th> <th>AK040521</th>	AK040521	AK040521
31	58.8	3.9	4845	11 <th>AK031490</th> <th>AK031490</th>	AK031490	AK031490
32	58.6	3.9	346	10 <th>BC057154</th> <th>BC057154</th>	BC057154	BC057154
33	57.4	3.8	724	14 <th>CD373179</th> <th>CD373179</th>	CD373179	CD373179
34	57.4	3.8	3590	13 <th>BU699337</th> <th>BU699337</th>	BU699337	BU699337
35	56.6	3.8	366	13 <th>BY395971</th> <th>BY395971</th>	BY395971	BY395971
36	56.6	3.8	784	14 <th>CA510759</th> <th>CA510759</th>	CA510759	CA510759
37	56.2	3.8	630	14 <th>CB577560</th> <th>CB577560</th>	CB577560	CB577560
38	55.8	3.7	408	9 <th>AL673391</th> <th>AL673391</th>	AL673391	AL673391
39	55.8	3.7	476	10 <th>BF376727</th> <th>BF376727</th>	BF376727	BF376727
40	55.8	3.7	593	10 <th>BG348697</th> <th>BG348697</th>	BG348697	BG348697
41	55.8	3.7	631	9 <th>AL640068</th> <th>AL640068</th>	AL640068	AL640068
42	55.8	3.7	804	9 <th>AL879839</th> <th>AL879839</th>	AL879839	AL879839
43	55.8	3.7	2048	11 <th>BC034364</th> <th>BC034364</th>	BC034364	BC034364
44	55.4	3.7	603	14 <th>CB581904</th> <th>CB581904</th>	CB581904	CB581904
45	55	3.7	379	9 <th>AM489725</th> <th>AM489725</th>	AM489725	AM489725

ALIGNMENTS

RESULT 1
AU204939
LOCUS
DEFINITION
AU204939 unpublished oligo-capped cDNA library, stage I4
Caenorhabditis elegans cDNA clone yk843e05 5', mRNA sequence.

ACCESSION
AU204939
VERSION
AU204939.1
KEYWORDS
GT:14836792
SOURCE
EST.
ORGANISM
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae

REFERENCE
1 (bases 1 to 653)
Kohara,Y., Shio-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
A complementary view of the C.elegans genome

JOURNAL
Unpublished
COMMENT
Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.

FEATURES
source
1..653
/organism="Caenorhabditis elegans"

```
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk843e05"
/sex="Hermaphrodite"
/rissue_type="whole animal"
/dev_stage="L4"
/clone_lib="unpublished oligo-capped cDNA library, stage
L4"

BASE COUNT      216 a      116 c      136 g      185 t
ORIGIN
Query Match      42.1%; Score 630.4; DB 9; Length 653;
Best Local Similarity 99.8%; Pred. No. 1.6e-116;
Matches 631; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 ATGCTGAAATTTGAAAATGTCAGAGCTAACAAAAATGAGCAAGAAACTGCATAG 60
      |||
      22 ATGCTGAAATTTGAAAATGTCAGAGCTAACAAAAATGAGCAAGAAACTGCATAG 81
OY      61 ACCACTTGTGGGAATCCTATTACATCAGTTGAGAAAGAAAACTTTTCATTCGA 120
      |||
      82 ACCACTTGTGGGAATCCTATTACATCAGTTGAGAAAGAAAACTTTTCATTCGA 141
OY      121 GTTGAAGCATTCATCGTACCTTAACAGTAAATTTTAAAGAAATGCGTGAAGAGAGTT 180
      |||
      142 GTTGAAGCATTCATCGTACCTTAACAGTAAATTTTAAAGAAATGCGTGAAGAGAGTT 201
OY      181 ATCTTCGAAACAGTGTCCATGATTATGACAAAGAACTGCGATTGATTCAGATCAGATGG 240
      |||
      202 ATCTTCGAAACAGTGTCCATGATTATGACAAAGAACTGCGATTGATTCAGATCAGATGG 261
OY      241 TTTGCAAGAAATGAAAAAGTTTGGGATGACAGAGTTGCTGCTCAAGTTTATCGAGCTGAC 300
      |||
      262 TTTGCAAGAAATGAAAAAGTTTGGGATGACAGAGTTGCTGCTCAAGTTTATCGAGCTGAC 321
OY      301 ACGAAATTTTGGCTCAATATTTTATCGAGCATATGTTGGTTGGCAACGCCCAATG 360
      |||
      322 ACGAAATTTTGGCTCAATATTTTATCGAGCATATGTTGGTTGGCAACGCCCAATG 381
OY      361 AGTGATCCCATATGATGATAAATGTTATGTCGCCGCCGTTGCAATCAAGAAATGAC 420
      |||
      382 AGTGATCCCATATGATGATAAATGTTATGTCGCCGCCGTTGCAATCAAGAAATGAC 441
OY      421 CAAAATGATATGATAAATGTTATGTAATATGATGATGAGCGAAATGTCGCCCAACT 480
      |||
      442 CAAAATGATATGATAAATGTTATGTAATATGATGATGAGCGAAATGTCGCCCAACT 501
OY      481 TCGCTGTCTCCAAAATTTGATGAAGGAGGCTCTCTCTTAAGCAAGCATGTTTCAAAGTT 540
      |||
      502 TCGCTGTCTCCAAAATTTGATGAAGGAGGCTCTCTCTTAAGCAAGCATGTTTCAAAGTT 561
OY      541 GGACAAAGCTTGAACCTTAAATTAATTCATTAATTCAGTAATTCAGGCGGAGCAAT 600
      |||
      562 GGACAAAGCTTGAACCTTAAATTAATTCATTAATTCAGTAATTCAGGCGGAGCAAT 621
OY      601 CAAGAAATATGTGAGCAGCAGATGAATGATATC 632
      |||
      622 CAAGAAATATGTGAGCAGCAGATGAATGATATC 653
DB
RESULT 2
BUI43186/ 768 bp mRNA linear EST 23-JAN-2002
LOCUS      BUI43186 unpublished oligo-capped cDNA library, C. elegans L1 stage
DEFINITION Caenorhabditis elegans cDNA clone yk1192h09 3', mRNA sequence.
ACCESSION      BUI43186
VERSION      BUI43186.1 GI:18303352
KEYWORDS      EST.
SOURCE      Caenorhabditis elegans
ORGANISM      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae
; Rhabditidae; Peloderinae; Caenorhabditis.
```

```
REFERENCE      1 (bases 1 to 768)
AUTHORS      Kohara,Y., Shih-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
TITLE      A complementary view of the C. elegans genome
JOURNAL      Unpublished
COMMENT      Contact: Tadasu Shin-i
      National Institute of Genetics
      111 Yata, Mishima, Shizuoka 411-8540, Japan
      Tel: 81-559-81-6856
      Fax: 81-559-81-6855
      Email: tshini@genes.nig.ac.jp.

FEATURES
source
      1..768
      Location/Qualifiers
      /organism="Caenorhabditis elegans"
      /mol_type="mRNA"
      /strain="N2"
      /db_xref="taxon:6239"
      /clone="yk1192h09"
      /sex="hermaphrodite"
      /rissue_type="whole animal"
      /dev_stage="L1"
      /clone_lib="unpublished oligo-capped cDNA library, C.
      elegans L1 stage"

BASE COUNT      232 a      151 c      146 g      237 t      2 others
ORIGIN
Query Match      40.0%; Score 599; DB 12; Length 768;
Best Local Similarity 96.8%; Pred. No. 3.4e-110;
Matches 630; Conservative 0; Mismatches 0; Indels 21; Gaps 1;

OY      847 AGATATGACTCAGACGACGTCACATTTGATGATTAAGCAAAAAGATCCATATGATCCCATG 906
      |||
      768 AGATATGACTCAGACGACGTCACATTTGATGATTAAGCAAAAAGATCCATATGATCCCATG 709
OY      907 ATTGGAGAAAAGTTAAGTTGGACAAAAGTTTACGTCATGCAAGCCCTGGCTCAGCAA 966
      |||
      708 ATTGGAGAAAAGTTAAGTTGGACAAAAGTTTACGTCATGCAAGCCCTGGCTCAGCAA 649
OY      967 TTCATAATACCTCCACGTCGCTTCATCTCAATTTTGGCAAAACCTGAAAGATATCTTAT 1026
      |||
      648 TTCATAATACCTCCACGTCGCTTCATCTCAATTTTGGCAAAACCTGAAAGATATCTTAT 589
OY      1027 GTGGGAATGATGTCAGATGACATCTTGAGACAGCTTTTCTATTCATATCAATTAATCA 1086
      |||
      588 GTGGGAATGATGTCAGATGACATCTTGAGACAGCTTTTCTATTCATATCAATTAATCA 529
OY      1087 TTTATGTCCTCCAGTTGTTATGCGGAAAAGTATTAATTTGSAACTGTTCCGCGCATGAG 1146
      |||
      528 TTTATGTCCTCCAGTTGTTATGCGGAAAAGTATTAATTTGSAACTGTTCCGCGCATGAG 469
OY      1147 TTCAAAAGACATTCAGATGAGATGAATTAATTTGAGAAAGATCTGCAAAAACCTCAACG 1206
      |||
      468 TTCAAAAGACATTCAGATGAGATGAATTAATTTGAGAAAGATCTGCAAAAACCTCAACG 409
OY      1207 CTGACTGTTCCAAAGCCCATGCTTCCCAAGAGATATGACAAATTAATTAAGTAAATCTG 1266
      |||
      408 CTGACTGTTCCAAAGCCCATGCTTCCCAAGAGATATGACAAATTAATTAAGTAAATCTG 359
OY      1267 ATTTCCAAAGGATGAGACTAGGCTGAGCTGATGATGATGATGATGATGATGATGATGAT 1326
      |||
      358 -----GATGAGACTAGGCTGAGCTGATGATGATGATGATGATGATGATGATGATGAT 310
OY      1327 TGTCAGCTACAGTGAATTCAGTTTCATGGAAGACTGATGAATTCATATTTGAGAGGCTGAG 1386
      |||
      309 TGTCAGCTACAGTGAATTCAGTTTCATGGAAGACTGATGAATTCATATTTGAGAGGCTGAG 250
OY      1387 GATGAAGATTTGATGAACCTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1446
      |||
      249 GATGAAGATTTGATGAACCTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 190
OY      1447 TGTCAGGCGCAGATTATGTTCTCAACCTCGAAAAAGTCAACTATATGA 1497
      |||
```

```

Db      189 TGTGAAGCGCAGATTATGTTCTACAACTCCGAAAAAGTACAACTATTGA 139
RESULT 3
LOCUS   BUI29361 761 bp mRNA linear EST 23-JAN-2002
DEFINITION BUI29361 unpublished oligo-capped cDNA library, C. elegans L1 stage
Caenorhabditis elegans cDNA clone yk1029h03 3', mRNA sequence.
ACCESSION BUI29361
VERSION   BUI29361.1 GI:18289518
KEYWORDS EST.
SOURCE   Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea
; Rhabditidae; Peleoderinae; Caenorhabditis.
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE    A complementary view of the C.elegans genome
JOURNAL  Unpublished
COMMENT  Contact: Tadasi Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
source   1..761
Location/Qualifiers
1..761
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/db_xref="taxon:6239"
/clone="yk1029h03"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, C.
elegans L1 stage"
BASE COUNT 233 a 150 c 144 g 232 t 2 others
ORIGIN
Query Match 39.5%; Score 591; DB 12; Length 761;
Best Local Similarity 96.6%; Pred. No. 1.4e-108;
Matches 622; Conservative 0; Mismatches 1; Indels 21; Gaps 1;
QY 854 ACTCAGACGACGTCACATTGATCATTCAGCAAAAGATCCATGATCCCATGATTGGA 913
|||||
Db 761 ACTCAGACGACGTCACATTGATCATTCAGCAAAAGATCCATGATCCCATGATTGGA 702
914 GAAAGTTAAGTTGGACAAAGTTGAGCTCATGCACCCCTGGCTCAGCATTCGAAT 973
|||||
Db 701 GAAAGTTAAGTTGGACAAAGTTGAGCTCATGCACCCCTGGCTCAGCATTCGAAT 642
974 ACCTCAGCTGCGCTTCGATTCGAAATTTGCAAAAGTGAAGATATCTTATTTGGGA 1033
|||||
Db 641 ACCTCAGCTGCGCTTCGATTCGAAATTTGCAAAAGTGAAGATATCTTATTTGGGA 582
1034 TGGATGGTCCAGATGCACATTGGAAGACAGTTTCTTCATCATATCAATTAATCTTATGT 1093
|||||
Db 581 TGGATGGTCCAGATGCACATTGGAAGACAGTTTCTTCATCATATCAATTAATCTTATGT 522
1094 TCCCACTTGGTTATGGGAAAAAGTATATTTGGAATGTTGCCGACAGATGATTTGAAG 1153
|||||
Db 521 TCCCACTTGGTTATGGGAAAAAGTATATTTGGAATGTTGCCGACAGATGATTTGAAG 462
1154 GAACATTGAGATGGATGATATCTTGGAGAAAGAAATCTGCAGAAACCTACCGCTTGACT 1213
461 GAACATTGAGATGGATGATATCTTGGAGAAAGAAATCTGCAGAAACCTACCGCTTGACT 402
1214 TGTTCAGGCAATGCCCTTCCCAAGAGAGATTAGACAAATTTAAGGTAATTCATTTCCA 1273
401 TGTTCAGGCAATGCCCTTCCCAAGAGAGATTAGACAAATTTTA----- 359

```

```

QY 1274 AACGGTAGACTACACCTTTGAAGCTGTCACATGTGTGMAAATCAGTTTATTTGCCAG 1333
|||||
Db 358 ----GGTAGACTACACCTTTGAAGCTGTCACATGTGTGMAAATCAGTTTATTTGCCAG 303
1334 CTACAGTGAATCAGTTTCATGGAAGACGTGAATTAATGTCATTTTCGACGCTGGGATGAG 1393
|||||
Db 302 CTACAGTGAATCAGTTTCATGGAAGACGTGAATTAATGTCATTTTCGACGCTGGGATGAG 243
QY 1394 AATTGATGAACTGTATGATGTGACTCCCATATATTTCTACCGATAGGATGTTGGAAG 1453
|||||
Db 242 AATTGATGAACTGTATGATGTGACTCCCATATATTTCTACCGATAGGATGTTGGAAG 183
QY 1454 CGCAGATTATGTTTTCACAACTCCGAAAAAGTACAACTATTGA 1497
182 CGCAGATTATGTTTTCACAACTCCGAAAAAGTACAACTATTGA 139
RESULT 4
LOCUS   CB404801 607 bp mRNA linear EST 15-MAY-2003
DEFINITION OSTR028D3_1 AD-wrmcDNA Caenorhabditis elegans cDNA, mRNA sequence.
ACCESSION CB404801
VERSION   CB404801.1 GI:30746528
KEYWORDS EST.
SOURCE   Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea
; Rhabditidae; Peleoderinae; Caenorhabditis.
AUTHORS 1 (bases 1 to 607)
Reboul,J., Vaglio,P., Rual,J.F., Jamesch,P., Martinez,M., Armstrong
,C.M., Li,S., Jacotot,L., Bertin,N., Janky,R., Moore,T., Hudson
,J.R., Harley,J.L., Brasch,M.A., Vandenhaute,J., Boulton,S.,
Endress,G.A., Jenna,S., Chevret,E., Papasotiropoulos,V., Tolias,P.P.,
Placek,J., Snyder,M., Huang,R., Chance,M.R., Lee,H.,
Doucette-Stamm,L., Hill,D.E. and Vidal,M.
C. elegans ORFome version 1.1: experimental verification of the
genome annotation and resource for proteome-scale protein
expression
Nac. Genet., (2003) In press
CONTACT: Vidal M
Marc Vidal Laboratory
Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739
Email: Marc.Vidal@dfci.harvard.edu
Sequence tag of gateway entry clones. The primers used were
designed on the predicted protein encoding ORF. C. elegans ORFome
cloning project : Contact david.hill@dfci.harvard.edu or
marc.vidal@dfci.harvard.edu
POLY-A-No.
FEATURES
source   1..607
Location/Qualifiers
1..607
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/db_xref="taxon:6239"
/strain="N2"
/sex="Hermaphrodite and male"
/tissue_type="whole animal"
/dev_stage="mixed stage"
/clone_lib="AD-wrmcDNA"
/clone="The AD-wrmcDNA library was generated with poly(A)+
RNA isolated from both hermaphrodite and male N2 worms of
all larval stages, embryos, adults and dauers and the
subsequent generation of cDNAs by poly(A) priming. The
cDNAs were cloned into pC86"
BASE COUNT 175 a 131 c 113 g 188 t
ORIGIN
Query Match 37.6%; Score 563.4; DB 14; Length 607;
Best Local Similarity 96.3%; Pred. No. 4.8e-103;
Matches 606; Conservative 0; Mismatches 1; Indels 22; Gaps 2;

```

```
QY      843  TCACAGATATGACTCAGACGACGTCCATTTGATCAATTTAGCAAAAGATCCAAATTTGATCC 902
      607  TCACAAGATATGACTCAGACGACGTCCATTTGATCAATTTAGCAAAAGATCCAAATTTGATCC 548
DB
QY      903  CATGATTTGGAGAAAAGTTAAAGTTGGACAAAAGTTAGCTCATGACCCCTTGGCTCA 962
DB      547  CATGATTTGGAGAAAAGTTAAAGTTGGACAAAAGTTAGCTCATGACCCCTTGGCTCA 488
QY      963  GCAATTTCAATTAACCTCCACGTCGCTTGCATTCCTCAATTTTGCAAAACCTGAAGATATCT 1022
DB      487  GCAATTTCAATTAACCTCCACGTCGCTTGCATTCCTCAATTTTGCAAAACCTGAAGATATCT 428
QY      1023  TATTTGGGAATGGATGGTCCAGATGACCTTGAAAGACAGTTTCCCTATTCATATCAATTA 1082
DB      427  TATTTGGGAATGGATGGTCCAGATGACCTTGAAAGACAGTTTCCCTATTCATATCAATTA 368
QY      1083  TACATTTATGTTCCAGAGTTGGTATGCGGAAAAGTATTAATTTGGAACTTTCCGCCAGA 1142
DB      367  TACATTTATGTTCCAGAGTTGGTATGCGGAAAAGTATTAATTTGGAACTTTCCGCCAGA 308
QY      1143  TGAGTTCAAGAACATTCAGATGGATGATTAATCTTGGAGAAAGATCTGCAGAAACCT 1202
DB      307  TGAGTTCAAGAACATTCAGATGGATGATTAATCTTGGAGAAAGATCTGCAGAAACCT 248
QY      1203  ACCGCTGACTGTTTCAAGCCAAATGCTTCCCAAGAGAGATTAGCAAAATTTAAGTAT 1262
DB      247  ACCGCTGACTGTTTCAAGCCAAATGCTTCCCAAGAGAGATTAGCAAAATTTA----- 194
QY      1263  TCTGATTTCCAAACGGGTAGACACTACGCTTGAAAGCTCTGACATGTTGAAAAACACTT 1322
DB      193  -----GTTAGGACATACGCTTGAAAGCTCTGACATGTTGAAAAACACTT 149
QY      1323  TATTTGCCAGCTACAGTGAATCAGTTTCATGAGAAAGTGAATTAATTCATATTTGACAG 1382
DB      148  TATTTGCCAGCTACAGTGAATCAGTTTCATGAGAAAGTGAATTAATTCATATTTGACAG 89
QY      1383  CTGGGATTAAGAAATTTGATGAATGATGATGATGAGTCCATGATTAATTCATACGATAG 1442
DB      88  CTGGGATTAAGAAATTTGATGAATGATGATGATGAGTCCATGATTAATTCATACGATAG 29
QY      1443  ATGTTGTGAGGCGACAGTTATGTTCTTAC 1471
DB      28  ATGTTGT- AAGCGCACAGTTATGTTCTTAC 1

RESULT 5
BU101701 581 bp mRNA linear EST 18-JAN-2002
LOCUS BU101701 unpublished oligo-capped cDNA library, C. elegans L1 stage
DEFINITION Caenorhabditis elegans cDNA clone yk1029h03 5', mRNA sequence.
ACCESSION BU101701
VERSION BU101701.1 GI:18244371
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
1 (bases 1 to 581)
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished
COMMENT Contact: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.
Location/Qualifiers
FEATURES
SOURCE 1..581
/oranism="Caenorhabditis elegans"
```

```
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1029h03"
/sex="hermaphrodite"
/russue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, C. elegans L1 stage"
BASE COUNT 189 a 103 c 123 g 166 t
ORIGIN
Query Match 37.3%, Score 559; DB 12; Length 581;
Best Local Similarity 100.0%; Pred. No. 3.6e-102;
Matches 559; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1  ATGCTGAATTTTCGAAAATTTGCAGAGCTTAACAAAATGCGACAGAAACCTGATTAAG 60
      23  ATGCTGAATTTTCGAAAATTTGCAGAGCTTAACAAAATGCGACAGAAACCTGATTAAG 82
DB
QY      61  ACCACTTTGTGGAAATCCTATTTCATCAGTTGAGAAAAGAAAACCTTTTCATTCCA 120
DB      83  ACCTAATTTGTGGAAATCCTATTTCATCAGTTGAGAAAAGAAAACCTTTTCATTCCA 142
QY      121  GTTGAGATTCATTCGTAACCTTACAGTTATTTTAACGAATGCGTGAAGAGAGT 180
DB      143  GTTGAGATTCATTCGTAACCTTACAGTTATTTTAACGAATGCGTGAAGAGAGT 202
QY      181  ATCTTGAACAGAGTGCATGATTAATGACAAGAACTGATTCGATTCAGATCAGATG 240
DB      203  ATCTTGAACAGAGTGCATGATTAATGACAAGAACTGATTCGATTCAGATCAGATG 262
QY      241  TTTGACAGAAATGAAAAGTTTGGGATACAGAGTTCTGCTCAGTTTATGAGCTGAC 300
DB      263  TTTGACAGAAATGAAAAGTTTGGGATACAGAGTTCTGCTCAGTTTATGAGCTGAC 322
QY      301  ACGAAATTTTGGCTCAATTTTATCGGACGATATGTTGGTTTGGCAAAAGCCGCAATG 360
DB      323  ACGAAATTTTGGCTCAATTTTATCGGACGATATGTTGGTTTGGCAAAAGCCGCAATG 382
QY      361  AGTATCCCAATTAATGATAATGATATGATGATGATGATGATGATGATGATGATGATG 420
DB      383  AGTATCCCAATTAATGATAATGATATGATGATGATGATGATGATGATGATGATGATG 442
QY      421  CAAAATGATATGATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
DB      443  CAAAATGATATGATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 502
QY      481  TCGCTGCTCCAAAATTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
DB      503  TCGCTGCTCCAAAATTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 562
QY      541  GGACAACTCTTGAACATAT 559
DB      563  GGACAACTCTTGAACATAT 581

RESULT 6
BU124772 557 bp mRNA linear EST 30-MAY-2003
LOCUS BU124772 unpublished oligo-capped cDNA library, C. elegans L1 stage
DEFINITION Caenorhabditis elegans cDNA clone yk1315b11 5', mRNA sequence.
ACCESSION BU124772
VERSION BU124772.2 GI:31246008
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
1 (bases 1 to 557)
TITLE A complementary view of the C.elegans genome
```

JOURNAL Unpublished
COMMENT On Jan 23, 2002 this sequence version replaced gi:18284913.
Contact: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.
Location/Qualifiers
1.557
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="YK1315b11"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, C.
elegans L1 stage"
BASE COUNT 180 a 99 c 118 g 158 t 2 others
ORIGIN
Query Match 34.8%; Score 521; DB 12; Length 557;
Best Local Similarity 99.8%; Pred. No. 1.6e-94;
Matches 532; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 3 GTCGTGAATTCGAAATTCGACAGCTAACAAAAATCGGACGAAACTCGATTAAGC 62
DB 25 GTCGTGAATTCGAAATTCGACAGCTAACAAAAATCGGACGAAACTCGATTAAGC 84
QY 63 CTACTGTGGGAATCCATATTTACATCAGTTCGAGAAAG-AAAACCTCTTCATTCGAG 121
DB 85 CTACTGTGGGAATCCATATTTACATCAGTTCGAGAAAGAAACCTCTTCATTCGAG 144
QY 122 TTGAACGATTCATCGTAACTTACGTTAATTTTAAAGAAATCGTGAAGGAAGAGTTA 181
DB 145 TTGAACGATTCATCGTAACTTACGTTAATTTTAAAGAAATCGTGAAGGAAGAGTTA 204
QY 182 TCTTGAACAGAGTCCATGATATGACAAAGACGCGATTCGATTCAGTCAAGTGGT 241
DB 205 TCTTGAACAGAGTCCATGATATGACAAAGACGCGATTCGATTCAGTCAAGTGGT 264
QY 242 TTGACGATTCGAAAGATTTGGGATGACAGATTCGCTGCTCAGTTTATCGAGCTGACA 301
DB 265 TTGACGATTCGAAAGATTTGGGATGACAGATTCGCTGCTCAGTTTATCGAGCTGACA 324
QY 302 CGAAATTTGGCTCAATATTTATCGGACGATATGTTGGTTGGCAACGCCGCAATGA 361
DB 325 CGAAATTTGGCTCAATATTTATCGGACGATATGTTGGTTGGCAACGCCGCAATGA 384
QY 362 GTGATCCCAATATGATTAATTTGATATGCTCCGCCGCTTGCATCAACGAAGAATACC 421
DB 385 GTGATCCCAATATGATTAATTTGATATGCTCCGCCGCTTGCATCAACGAAGAATACC 444
QY 422 AAAATGATATGTAATATGTAATATTTGATGATGCGGAAATCGTCGGCAAACTT 481
DB 445 AAAATGATATGTAATATGTAATATTTGATGATGCGGAAATCGTCGGCAAACTT 504
QY 482 CGCTGTCTCAAAATTCGATGAAGGAGGCTCTCCCTAAGCAAGCATCGTTTC 534
DB 505 CGCTGTCTCAAAATTCGATGAAGGAGGCTCTCCCTAAGCAAGCATCGTTTC 557
RESULT 7
AU216908/c 677 bp mRNA linear EST 17-JUL-2001
LOCUS AU216908 unpublished oligo-capped cDNA library, stage I4
DEFINITION Caenorhabditis elegans cDNA clone yk843e05 3', mRNA sequence.
ACCESSION AU216908
VERSION AU216908.1 GI:14855065
KEYWORDS EST.

SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE 1 (bases 1 to 677)
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers
1.677
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="YK843e05"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L4"
/clone_lib="unpublished oligo-capped cDNA library, stage
L4"
BASE COUNT 210 a 136 c 130 g 201 t
ORIGIN
Query Match 34.7%; Score 519; DB 9; Length 677;
Best Local Similarity 96.3%; Pred. No. 4e-94;
Matches 550; Conservative 0; Mismatches 0; Indels 21; Gaps 1;
QY 927 TGGACAAAGTTTGACCTCATCGACCCCTTGCTCAGCAATTCATTAACCTCCAGCTGC 986
DB 677 TGGACAAAGTTTGACCTCATCGACCCCTTGCTCAGCAATTCATTAACCTCCAGCTGC 618
QY 987 TTGCATTCGAAATTTGCAAAACGGAAGATATCTATTGGGGAATGAGATGGTCACA 1046
DB 617 TTGCATTCGAAATTTGCAAAACGGAAGATATCTATTGGGGAATGAGATGGTCACA 558
QY 1047 TGCACCTTGAAGACAGTTTCTTATTCATATCAATATTAACATTTATGTTCCAGTTGGTTA 1106
DB 557 TGCACCTTGAAGACAGTTTCTTATTCATATCAATATTAACATTTATGTTCCAGTTGGTTA 498
QY 1107 TCGGAAAGATATATTTGGAACCTGTCGCCGACAGTGAAGTCAAGAACATTCAGATG 1166
DB 497 TCGGAAAGATATATTTGGAACCTGTCGCCGACAGTGAAGTCAAGAACATTCAGATG 438
QY 1167 GGATGAACTCTGGACAAAGAAATTCGACAAACCCATACCGCTTGACTTGTCAAGCCAT 1226
DB 437 GGATGAACTCTGGACAAAGAAATTCGACAAACCCATACCGCTTGACTTGTCAAGCCAT 378
QY 1227 GCGTTCGCAAGAGAGATTAGACAAATTTAAGTAATTTGATTTCCAAACGGGTAGAGCT 1286
DB 377 GCGTTCGCAAGAGAGATTAGACAAATTTA-----GGTAGACT 339
QY 1287 ACGCTTGAAGCTGTGACATGTGGAATAATCAATTTATTTGTCCAGCTACAGTGAATTC 1346
DB 338 ACGCTTGAAGCTGTGACATGTGGAATAATCAATTTATTTGTCCAGCTACAGTGAATTC 279
QY 1347 AGTTCATGAACATCATTAATGTCATTTGCAAGCGCTGGAGTGAAGAAATTTGATCAACT 1406
DB 278 AGTTCATGAACATCATTAATGTCATTTGCAAGCGCTGGAGTGAAGAAATTTGATCAACT 219
QY 1407 GTATGATGTGACCTCCATGATATTTCTACCGATAGGATGGTGAAGCCACAGTTATGT 1466
DB 218 GTATGATGTGACCTCCATGATATTTCTACCGATAGGATGGTGAAGCCACAGTTATGT 159
QY 1467 TCTACACCTCGAAAGATACACTATTGA 1497
|||||

Db 158 TCACACCTCCGAAAAAGTACACTATGCA 128

RESULT 8
Bj153087/c
LOCUS
DEFINITION Bj153087 unpublished oligo-capped cDNA library, C. elegans L1 stage
Caenorhabditis elegans cDNA clone YK1315b11 3', mRNA sequence.
ACCESSION Bj153087
VERSION Bj153087.1 GI:18321072
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea
; Rhabdilitidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 686)
AUTHORS Kohara,Y., Shih-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.,
and Sugano,S.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished
COMMENT Contact: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.
Location/Qualifiers
1. 686
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="YK1315b11"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, C.
elegans L1 stage"

BASE COUNT 214 a 133 c 131 g 204 t 4 others

ORIGIN

Query Match 34.1%; Score 510; DB 12; Length 686;
Best Local Similarity 96.1%; Pred. No. 2.6e-92;
Matches 541; Conservative 0; Mismatches 1; Indels 21; Gaps 1;

QY 935 AGTTGAGCTCATGAGCCCTTGGCTCAGCAATTCATACCTCCAGCTCGCTTCGATTC 994
|||||
Db 686 AGTTGAGCTCATGAGCCCTTGGCTCAGCAATTCATACCTCCAGCTCGCTTCGATTC 627
|||||

QY 995 TCAATATTTGCAAACTGAGAGATATCTTATTGGGGAATGAGATGCTCAGATGCACCTTG 1054
|||||
Db 626 TCAATATTTGCAAACTGAGAGATATCTTATTGGGGAATGAGATGCTCAGATGCACCTTG 567
|||||

QY 1055 AAGCAGGTTTCCTATTCATATCATATATACATTATGTCCTCCAGTTGGTATGGGAAA 1114
|||||
Db 566 AAGCAGGTTTCCTATTCATATCATATATACATTATGTCCTCCAGTTGGTATGGGAAA 507
|||||

QY 1115 AGTAATATTTGGAAGCTTGTCCGCGAGATGATTCAAAGAAACATTCAGATGGATGAT 1174
|||||
Db 506 AGTAATATTTGGAAGCTTGTCCGCGAGATGATTCAAAGAAACATTCAGATGGATGAT 447
|||||

QY 1175 ACTTGGAAGAAAGATCTGAGAAACCTTACCGCTTGACTGTTCAAGCCATGCTTCCC 1234
|||||
Db 446 ACTTGGAAGAAAGATCTGAGAAACCTTACCGCTTGACTGTTCAAGCCATGCTTCCC 387
|||||

QY 1235 AAGGAGATTTGACAAATTTAAGTAATTGATTTCCAAAGGGGTAGAGACTAGCCCTTG 1294
|||||
Db 386 AAGGAGATTTGACAAATTTA-----GGTAGACTAGCCCTTG 348
|||||

QY 1295 AAGCTGCTGACATGTGTGAAATCAGTTATTTGTCCAGCTACAGATGAATCAGTTCAATG 1354
|||||
Db 347 AAGCTGCTGACATGTGTGAAATCAGTTATTTGTCCAGCTACAGATGAATCAGTTCAATG 288
|||||

QY 1355 GAAGACTGATTAATGTCATATTTGACGCGCTGGATGAGAATTTGATGAACGTGATGANG 1414
|||||
Db 287 GAAGACTGATTAATGTCATATTTGACGCGCTGGATGAGAATTTGATGAACGTGATGANG 228
|||||

QY 1415 TGGACTCCCATGATATCTTACCGATAGAGATGCTGTGAAGCGCACAGTTATGTTCTACAC 1474
|||||
Db 227 TGGACTCCCATGATATCTTACCGATAGAGATGCTGTGAAGCGCACAGTTATGTTCTACAC 168
|||||

QY 1475 CTCGAAAAAGTACACTATGCA 1497
|||||
Db 167 CTCGAAAAAGTACACTATGCA 145
|||||

RESULT 9
Bj115564
LOCUS
DEFINITION Bj115564 unpublished oligo-capped cDNA library, C. elegans L1 stage
Caenorhabditis elegans cDNA clone YK1192h09 5', mRNA sequence.
ACCESSION Bj115564
VERSION Bj115564.1 GI:18275658
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea
; Rhabdilitidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 519)
AUTHORS Kohara,Y., Shih-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.,
and Sugano,S.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished
COMMENT Contact: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.
Location/Qualifiers
1. 519
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="YK1192h09"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, C.
elegans L1 stage"

BASE COUNT 170 a 91 c 109 g 149 t

ORIGIN

Query Match 33.6%; Score 502.4; DB 12; Length 519;
Best Local Similarity 99.8%; Pred. No. 8.5e-91;
Matches 503; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTGATATTTTCGAAAATTTGTCAAGAGCTTAACAAAAATGCGACAGAAAACTGGATAG 60
|||||
Db 16 ATGCTGATATTTTCGAAAATTTGTCAAGAGCTTAACAAAAATGCGACAGAAAACTGGATAG 75
|||||

QY 61 ACCTACTGTGGGAATCCTATTTACATCAGTTCGAGAAAGAAAACTTCTTCATTCCA 120
|||||
Db 76 ACCTACTGTGGGAATCCTATTTACATCAGTTCGAGAAAGAAAACTTCTTCATTCCA 135
|||||

QY 121 GTTGAAGCATTCATCGTAACCTTAACGTTAATTTTAAAGATCGGTGAAGAGAGACTT 180
|||||
Db 136 GTTGAAGCATTCATCGTAACCTTAACGTTAATTTTAAAGATCGGTGAAGAGAGAGACTT 195
|||||

QY 181 ATCTTCGAAAGAGTGTCTCATGATTTATGACAAGAACTCGATTCGATTCAGTCAGATNG 240
|||||
Db 196 ATCTTCGAAAGAGTGTCTCATGATTTATGACAAGAACTCGATTCGATTCAGTCAGATNG 255
|||||

Qy	241	TTTCACGAATGTGAAAAAGTTTGGCGATACAGAGTTCTGCGTCAAGTTTATGCGAGCTAC	300
Db	256	TTTTCACGAATGTGAAAAAGTTTGGCGATACAGAGTTCTGCGTCAAGTTTATGCGAGCTAC	315
Qy	301	ACGAATTTTGGCTCAATATTTTATGCGAGCATATGTTGGTTTGGCAACGCCGCATG	360
Db	316	ACGAATTTTGGCTCAATATTTTATGCGAGCATATGTTGGTTTGGCAACGCCGCATG	375
Qy	361	AGTATATCCCATATNGGATMAAATTGTATATGCTCCGCCCTTGCATCAACGAAGATAC	420
Db	376	AGTATATCCCATATNGGATMAAATTGTATATGCTCCGCCCTTGCATCAACGAAGAAATAC	435
Qy	421	CAAAATGATATGCTAAATTTATGTAAATTAATTCGATTGATGCGCAATCGTCGCCCAACT	480
Db	436	CAAAATGATATGCTAAATTTATGTAAATTAATTCGATTGATGCGCAATCGTCGCCCAACT	495
Qy	481	TCGCTGTCTCAAAATTCGATGAA	504
Db	496	TCGCTGTCTCAAAATTCGATGAA	519

RESULT	10
LOCUS	AV197399 380 bp mRNA linear EST 26-JUL-1999
DEFINITION	AV197399 Yuiji Kohara unpublished cDNA:Strain N2 hermaphrodite embryo Caenorhabditis elegans clone yk65h11 5', mRNA sequence.
ACCESSION	AV197399
VERSION	AV197399.1 GI:5581170
KEYWORDS	EST.
SOURCE	Caenorhabditis elegans
ORGANISM	Caenorhabditis elegans
REFERENCE	Eukaryota; Metazoa; Nematozoa; Chromadorea; Rhabdilitida; Rhabditoidea; Rhabdilitidae; Poloderinae; Caenorhabditis.
AUTHORS	Kohara,Y., Shio-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Mitsuki,H., Nishiyaki,A., Motokashii,T., Zeng,Q., Watanabe,H., Sugimoto,A., Sano,M., Miyata,K., Mitani,Y., Iida,K., Uesugi,H., Sugiyama,Y. and Nomoto,H.
TITLE	Expressed genes in C.elegans
JOURNAL	Unpublished
COMMENT	Contact: Yuiji Kohara Genome Biology Lab. National Institute of Genetics Yata 1111, Mishima, Shizuoka 411, Japan Tel: 81-559-81-6854 Fax: 81-559-81-6855 Email: ykohara@lab.nig.ac.jp.
FEATURES	Location/Qualifiers
source	1..380
	/organism="Caenorhabditis elegans"
	/mol_type="mRNA"
	/strain="N2"
	/db_xref="taxon:6239"
	/clone="yk65h11"
	/sex="hermaphrodite"
	/dex_stage="embryo"
	/clone_lib="Yuiji Kohara unpublished cDNA:Strain N2 hermaphrodite embryo"
BASE COUNT	125 a 63 c 83 g 108 t 1 others
ORIGIN	
Query Match	24.4%; Score 366; DB 9; Length 380;
Best Local Similarity	99.7%; Pred. No. 2.2e-63;
Matches 377; Conservative	0; Mismatches 0; Indels 1; Gaps 1;
QY	16 AAAATGTCGAGTGAACAATAATCGACGAGAATCGATTAGACTACTGTG -GGGA 74
D0	3 AAAATGTCGAGTGAACAATAATCGACGAGAATCGATTAGACTACTGTG -GGGA 62
OY	75 ATCCTATTTCACATGATTCGAGAAAGAAAAAATCTTTCAATCCAGTTGAAGCATTCNA 134
D0	63 ATCCTATTTCACATGATTCGAGAAAGAAAAAATCTTTCAATCCAGTTGAAGCATTCNA 122

QY	135	UCGTAACCTTCAGATTATTTTAAAGCATGCCTGAAGAGAGATTACTTCGAACAGT	194
Db	123	TCGTACCTTCAGATTAAATTTTAAAGCATGCCTGAAGAGAGATTACTTCGAACAGT	182
QY	195	GGTCGATGATTATGACAGAAGCTCGGATTCGATTCAATCAATGCTTGGACGAATTGA	254
Db	183	GGTCGATGATTATGACAGAAGCTCGGATTCGATTCAATCAATGCTTGGACGAATTGA	242
QY	255	AAAAGTTTGGCGATACAGAGTCTCGGCTCAGTTTATCGAGCTGACACGAATTTTGGCT	314
Db	243	AAAAGTTTGGCGATACAGAGTCTCGGCTCAGTTTATCGAGCTGACACGAATTTTGGCT	302
QY	315	CAATTTTATTCGACGATATGTTGGTTTGGCAACGCCGCAATGATGCCAATAT	374
Db	303	CAATTTTATTCGACGATATGTTGGTTTGGCAACGCCGCAATGATGCCAATAT	362
QY	375	GGATTAATAATTGATATGC	392
Db	363	GGATTAATAATTGATATGC	380

[illegible]

```
|||||
Db 121 TACAGTATATTTTAAAGATGCGTGAAGAGAGATTTCTTCCAAACAGTGGTCATGA 180
Oy 204 TTATGACAAAGTGCATGATTCAGATGAGTGGTTCAGCAATTTGAAAAAGTTTG 263
Db 181 TTATGACAAAGTGCATGATTCAGATGAGTGGTTCAGCAATTTGAAAAAGTTTG 240
Oy 264 CGATATACAGATCTGCTCAGTTTATCGAGCTGACAGAAATTTTGGCTCAATATTTT 323
Db 241 CGGATACAGATCTGCTCAGTTTATCGAGCTGACAGAAATTTTGGCTCAATATTTT 300
Oy 324 ATCGAGCATATGTTTGGTTGGCAACGCCGCAATGTCATCCCATATGATTAAT 383
Db 301 ATCGAGCATATGTTTGGTTGGCAACGCCGCAATGATGATCCCATATGATTAAT 360

RESULT 12
LOCUS AU111588 340 bp mRNA linear EST 19-OCT-2000
DEFINITION AU111588 unpublished oligo-capped cDNA library Caenorhabditis
ACCESSION AU111588
VERSION AU111588
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
AUTHORS ; Rhabditidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 340)
Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C. elegans genome
JOURNAL Unpublished
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
FEATURES
source location/Qualifiers
1..340
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk732e5"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="varied"
/clone_lib="unpublished oligo-capped cDNA library"
BASE COUNT 111 a 58 c 72 g 98 t 1 others
ORIGIN
Query Match 20.6%; Score 308; DB 9; Length 340;
Best Local Similarity 100.0%; Pred. No. 9.7e-52;
Matches 308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 GTCTGAATTTCTGAAATTTGTCAGAGCTAACAAAAATCGACAGAAAACCTCGATAAGC 62
Db 33 GTCTGAATTTCTGAAATTTGTCAGAGCTAACAAAAATCGACAGAAAACCTCGATAAGC 92
Oy 63 CTACTTGTGGGAATCTTATTTACATCAGTTCGAGAGAAAAAAGCTTTTCATTTCCAGT 122
Db 93 CTACTTGTGGGAATCTTATTTACATCAGTTCGAGAGAAAAAAGCTTTTCATTTCCAGT 152
Oy 123 TGAAGCATTCATCTTAACCTTACAGTTAATTTTACCAATGCCGGAAGAGAGGTTAT 182
Db 153 TGAAGCATTCATCTTAACCTTACAGTTAATTTTACCAATGCCGGAAGAGAGGTTAT 212
Oy 183 CTTCGAACAGTGTGTCATGATTTATGACAGAACTGCAATTCGATTCAGATGATGTT 242
Db 213 CTTCGAACAGTGTGTCATGATTTATGACAGAACTGCAATTCGATTCAGATGATGTT 272
```

```
Oy 243 TGCACGAATTCGAAAGATTTTCGCATACAGAGTTCTGCTCAGTTTATCGAGCTGACAC 302
Db 273 TGCACGAATTCGAAAGATTTTCGCATACAGAGTTCTGCTCAGTTTATCGAGCTGACAC 332
Oy 303 GAAATTTT 310
Db 333 GAAATTTT 340

RESULT 13
LOCUS AU116296/c 300 bp mRNA linear EST 19-OCT-2000
DEFINITION AU116296 unpublished oligo-capped cDNA library Caenorhabditis
ACCESSION AU116296
VERSION AU116296
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
AUTHORS ; Rhabditidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 300)
Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C. elegans genome
JOURNAL Unpublished
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
FEATURES
source location/Qualifiers
1..300
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk743a4"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="varied"
/clone_lib="unpublished oligo-capped cDNA library"
BASE COUNT 95 a 63 c 54 g 88 t
ORIGIN
Query Match 14.7%; Score 220; DB 9; Length 300;
Best Local Similarity 100.0%; Pred. No. 4.6e-34;
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1278 GTAGAGACTACGCCCTTGAAGCTGCTGACATGTGTAATAATCAGTTTATTTCTCCAGCTAC 1337
Db 300 GTAGAGACTACGCCCTTGAAGCTGCTGACATGTGTAATAATCAGTTTATTTCTCCAGCTAC 241
Oy 1338 AGTGAATTCAGTTCTATGGAACACTGATTAATGTCAATTTCCAGCGCTGGATGAAGATT 1397
Db 240 AGTGAATTCAGTTCTATGGAACACTGATTAATGTCAATTTCCAGCGCTGGATGAAGATT 181
Oy 1398 TGATGAACCTGATGATGTGGAGCTCCCATGATTTCTACCGATAGAGATGGTGAAGCGCA 1457
Db 180 TGATGAACCTGATGATGTGGAGCTCCCATGATTTCTACCGATAGAGATGGTGAAGCGCA 121
Oy 1458 CAGTTATGTCTACAACTCCGAAAAAGTACACTATTGA 1497
Db 120 CAGTTATGTCTACAACTCCGAAAAAGTACACTATTGA 81

RESULT 14
LOCUS AU115428 300 bp mRNA linear EST 19-OCT-2000
DEFINITION AU115428 unpublished oligo-capped cDNA library Caenorhabditis
```


ACCESSION AUI15428
VERSION AUI15428.1 GI:10928995
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Eukaryota; Metazoa; Nematoda; Chromodorea; Rhabdilitida; Rhabditoidea
AUTHORS ; Rhabdilitidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 300)
Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C. elegans genome
JOURNAL Unpublished
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.

FEATURES
source
1..300
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk732e5"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="varied"
/clone_lib="unpublished oligo-capped cDNA library"

BASE COUNT 94 a 60 c 54 g 89 t 3 others
ORIGIN

Query Match 14.4%; Score 215; DB 9; Length 300;
Best Local Similarity 100.0%; Pred. No. 4.7e-33;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1283 GACTACGCTTGAGCTGCTACATGCTGAAATTCATTTGTCACCTACAGTCA 1342
|||
DB 300 GACTACGCTTGAGCTGCTACATGCTGAAATTCATTTGTCACCTACAGTCA 241
|||
OY 1343 AATCAGTTCAGGAAGACTGATTAATGTCATTTGCGAGCTGGATGAGAATTTGATG 1402
|||
DB 240 AATCAGTTCAGGAAGACTGATTAATGTCATTTGCGAGCTGGATGAGAATTTGATG 181
|||
OY 1403 AACTGATGATGCTGCTCCATGATATTTACCGATAGATGCTGAGCGACAGT 1462
|||
DB 180 AACTGATGATGCTGCTCCATGATATTTACCGATAGATGCTGAGCGACAGT 121
|||
OY 1463 ATGTTCTACACCTCCGAAAAGTACAACTATTGA 1497
|||
DB 120 ATGTTCTACACCTCCGAAAAGTACAACTATTGA 86
|||

RESULT 15
AUI12450 337 bp mRNA linear EST 30-MAY-2003
LOCUS AUI12450 unpublished oligo-capped cDNA library Caenorhabditis
DEFINITION AUI12450
ACCESSION AUI12450
VERSION AUI12450.2 GI:31237498
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Eukaryota; Metazoa; Nematoda; Chromodorea; Rhabdilitida; Rhabditoidea
AUTHORS ; Rhabdilitidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 337)
Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C. elegans genome
JOURNAL Unpublished
COMMENT On Oct 19, 2000 this sequence version replaced gi:10926017.

Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.

FEATURES
source
1..337
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk743a4"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="varied"
/clone_lib="unpublished oligo-capped cDNA library"

BASE COUNT 107 a 59 c 74 g 92 t 5 others
ORIGIN

Query Match 12.4%; Score 186.2; DB 9; Length 337;
Best Local Similarity 97.9%; Pred. No. 2.9e-27;
Matches 188; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 ATGCTGAATTTCTGAAATTTGTCAGAGCTAACAAAAATCGACAGAAAATCGATPAG 60
|||
DB 23 ATGCTGAATTTCTGAAATTTGTCAGAGCTAACAAAAATCGACAGAAAATCGATPAG 82
|||
OY 61 ACCTACTGTGGAATCCTATTATTCATCAGTTCGAGAAAGAAAACCTTCTTCATTCCA 120
|||
DB 83 ACCTACTGTGGAATCCTATTATTCATCAGTTCGAGAAAGAAAACCTTCTTCATTCCA 142
|||
OY 121 GTTGAAGCATTCATTCGTAACCTTACAGTTAATTTTAACGAATGCGTAAGAGAGATT 180
|||
DB 143 GTTGAAGCATTCATTCGTAACCTTACAGTTAATTTTAACGAATGCGTAAGAGAGATT 202
|||
OY 181 ATCTTCGAAACA 192
|||
DB 203 ATCTTCGAAACA 214
|||

Search completed: September 2, 2003, 17:41:43
Job time : 3449 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2003, 17:41:55 ; Search time 3440 Seconds
(without alignments)
10576.686 Million cell updates/sec

Title: US-09-872-523-6
Perfect score: 1497
Sequence: 1 atgtctgaattcttgaataa.....cgaaaagtaacatttga 1497

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 0

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: em_estha:*
2: em_esthum:*
3: em_estin:*
4: em_estimu:*
5: em_estrov:*
6: em_estropl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pin:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	581	38.8	653	9 AU204939	AU204939 AU204939
2	559	37.3	581	12 BJ101701	BJ101701 BJ101701
3	453	30.3	519	12 BJ115564	BJ115564 BJ115564
4	433	28.9	557	12 BJ124772	BJ124772 BJ124772

c 5	414	27.7	768	12 BJ143186
c 6	385	25.7	607	14 CB404801
c 7	360	24.0	360	14 CA4305
c 8	356	23.8	761	12 BJ129361
c 9	334	22.3	677	9 AU216908
c 10	322	21.5	380	9 AV197359
c 11	308	20.6	340	9 AU111588
c 12	306	20.4	686	12 BJ153087
c 13	220	14.7	300	9 AU116296
c 14	215	14.4	300	9 AU115428
c 15	186	12.4	337	9 AU112450
c 16	66	4.4	505	14 CB385690
c 17	23	1.5	1291	10 BG503414
c 18	22	1.5	265	12 BM632811
c 19	22	1.5	304	14 CB698306
c 20	22	1.5	426	14 CB695113
c 21	22	1.5	507	28 AZ907625
c 22	22	1.5	555	14 CB545230
c 23	21	1.4	335	14 H34557
c 24	21	1.4	399	28 BH787508
c 25	21	1.4	453	14 CB395614
c 26	21	1.4	453	14 CB395661
c 27	21	1.4	467	10 BE416260
c 28	21	1.4	493	29 CC063342
c 29	21	1.4	635	12 BJ290111
c 30	21	1.4	638	29 B2724245
c 31	21	1.4	758	28 A2124760
c 32	21	1.4	776	14 CB963916
c 33	21	1.4	839	13 BU358433
c 34	21	1.4	872	9 AU121368
c 35	21	1.4	1032	29 B2824126
c 36	20	1.3	283	10 BF525236
c 37	20	1.3	333	28 BH555727
c 38	20	1.3	360	9 AV194140
c 39	20	1.3	360	14 C45578
c 40	20	1.3	360	14 D74616
c 41	20	1.3	385	9 AM640232
c 42	20	1.3	404	29 B2491812
c 43	20	1.3	405	14 T94263
c 44	20	1.3	420	14 CB798312
c 45	20	1.3	430	13 BU804167

ALIGNMENTS

RESULT 1
AU204939
LOCUS
DEFINITION
AU204939 unpublished oligo-capped cDNA library, stage L4
Caenorhabditis elegans cDNA clone yk843e05 5', mRNA sequence.
ACCESSION
AU204939
VERSION
AU204939.1 GI:114836792
KEYWORDS
EST.
SOURCE
Caenorhabditis elegans
ORGANISM
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae
; Rhabditidae; Pelodderinae; Caenorhabditis.
REFERENCE
1 (bases 1 to 653)
Kohara,Y., Shih-I.T., Thlerry-Mleg,J., Suzuki,Y.
and Sugano,S.
A complementary view of the C.elegans genome
TITLE
JOURNAL
COMMENT
Unpublished
Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers
1. 653
/organism="Caenorhabditis elegans"

```
/mol_type="mrna"
/db_xref="taxon:6239"
/clone="YK843e05"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L4"
/clone_lib="unpublished oligo-capped cdna library, stage
L4"

BASE COUNT      216 a      116 c      136 g      185 t

Query Match      38.8%; Score 581; DB 9; Length 653;
Best Local Similarity 99.8%; Pred. No. 9.7e-303;
Matches 631; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ATGCTGAATTTTCGAAATTTGTCAGAGCTACACAAAATCGACAGAAAACCTCGATAG 60
    |||||||
DB 22 ATGCTGAATTTTCGAAATTTGTCAGAGCTACACAAAATCGACAGAAAACCTCGATAG 81
    |||||||

OY 61 ACCTACTTGTGGGAATCCTATTACATCAGTTGCGAAGAAAACCTTCTTTCATTC 120
    |||||||
DB 82 ACCTACTTGTGGGAATCCTATTACATCAGTTGCGAAGAAAACCTTCTTTCATTC 141
    |||||||

OY 121 GTTGAACGATTCATCGTAACCTTACGATTAATTTTAACGAAATCCGTAAGAGAGACTT 180
    |||||||
DB 142 GTTGAACGATTCATCGTAACCTTACGATTAATTTTAACGAAATCCGTAAGAGAGACTT 201
    |||||||

OY 181 ATCTTGAACAGAGGTCATGATTATGACAAAGATCGATTCGATTCAGATCAGATGG 240
    |||||||
DB 202 ATCTTGAACAGAGGTCATGATTATGACAAAGATCGATTCGATTCAGATCAGATGG 261
    |||||||

OY 241 TTTGCAGAAATGAAAAAGTTTGGGATACAGAGTTCTGCTCAGTTTATCGAGCTGAC 300
    |||||||
DB 262 TTTGCAGAAATGAAAAAGTTTGGGATACAGAGTTCTGCTCAGTTTATCGAGCTGAC 321
    |||||||

OY 301 ACGAAATTTTGGCCATATTTTATCGAGCATATGTTGGTTTGGCAAGAGCGGCATG 360
    |||||||
DB 322 ACGAAATTTTGGCCATATTTTATCGAGCATATGTTGGTTTGGCAAGAGCGGCATG 381
    |||||||

OY 361 AGTATCCCATATGATGATAAATTTGATATGCTCCGCCGCTTGCATCAACGAAGAATAC 420
    |||||||
DB 382 AGTATCCCATATGATGATAAATTTGATATGCTCCGCCGCTTGCATCAACGAAGAATAC 441
    |||||||

OY 421 CAAAATGATATGTAATTTATGTAATATTCATGATGCGGAATTCGTGCGCAACT 480
    |||||||
DB 442 CAAAATGATATGTAATTTATGTAATATTCATGATGCGGAATTCGTGCGCAACT 501
    |||||||

OY 481 TCGCTGTCTCCAAATTCGATGAGGAGGCTCTCCTAAGCAAGCATCGTTCAAAAGTT 540
    |||||||
DB 502 TCGCTGTCTCCAAATTCGATGAGGAGGCTCTCCTAAGCAAGCATCGTTCAAAAGTT 561
    |||||||

OY 541 GGACAAGCTCTTGAACATTTTAAATTTTCAATTTCTACTGAATACGGGTAGCGCAAT 600
    |||||||
DB 562 GGACAAGCTCTTGAACATTTTAAATTTTCTACTGAATACGGGTAGCGGTAGCTT 621
    |||||||

OY 601 CAAGAATATGTGACGACGAGATGATATC 632
    |||||||
DB 622 CAAGAATATGTGACGACGAGATGATATC 653
    |||||||

RESULT 2
BUI01701      581 bp      mRNA      linear      EST 18-JAN-2002
LOCUS      BUI01701 unpublished oligo-capped cdna library, C. elegans L1 stage
DEFINITION      Caenorhabditis elegans cdna clone yk1029h03 5', mRNA sequence.
ACCESSION      BUI01701
VERSION      BUI01701.1 GI:18244371
KEYWORDS      EST.
SOURCE      Caenorhabditis elegans
ORGANISM      Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
; Rhabditidae; Peloderinae; Caenorhabditis.
```

```
REFERENCE      1 (bases 1 to 581)
AUTHORS      Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE      A complementary view of the C.elegans genome
JOURNAL      Unpublished
COMMENT      Contact: Tadasi Shin-I
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.
Location/Qualifiers
1..581
/organism="Caenorhabditis elegans"
/mol_type="mrna"
/db_xref="taxon:6239"
/clone="YK1029h03"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cdna library, C.
elegans L1 stage"

BASE COUNT      189 a      103 c      123 g      166 t

Query Match      37.3%; Score 559; DB 12; Length 581;
Best Local Similarity 100.0%; Pred. No. 7.7e-291;
Matches 559; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGCTGAATTTTCGAAATTTGTCAGAGCTACACAAAATCGACAGAAAACCTCGATAG 60
    |||||||
DB 23 ATGCTGAATTTTCGAAATTTGTCAGAGCTACACAAAATCGACAGAAAACCTCGATAG 82
    |||||||

OY 61 ACCTACTTGTGGGAATCCTATTACATCAGTTGCGAAGAAAACCTTCTTTCATTC 120
    |||||||
DB 83 ACCTACTTGTGGGAATCCTATTACATCAGTTGCGAAGAAAACCTTCTTTCATTC 142
    |||||||

OY 121 GTTGAACGATTCATCGTAACCTTACGATTAATTTTAACGAAATCCGTAAGAGAGACTT 180
    |||||||
DB 143 GTTGAACGATTCATCGTAACCTTACGATTAATTTTAACGAAATCCGTAAGAGAGACTT 202
    |||||||

OY 181 ATCTTGAACAGAGGTCATGATTATGACAAAGATCGATTCGATTCAGATCAGATGG 240
    |||||||
DB 203 ATCTTGAACAGAGGTCATGATTATGACAAAGATCGATTCGATTCAGATCAGATGG 262
    |||||||

OY 241 TTTGCAGAAATGAAAAAGTTTGGGATACAGAGTTCTGCTCAGTTTATCGAGCTGAC 300
    |||||||
DB 263 TTTGCAGAAATGAAAAAGTTTGGGATACAGAGTTCTGCTCAGTTTATCGAGCTGAC 322
    |||||||

OY 301 ACGAAATTTTGGCCATATTTTATCGAGCATATGTTGGTTTGGCAAGAGCGGCATG 360
    |||||||
DB 323 ACGAAATTTTGGCCATATTTTATCGAGCATATGTTGGTTTGGCAAGAGCGGCATG 382
    |||||||

OY 361 AGTATCCCATATGATGATAAATTTGATATGCTCCGCCGCTTGCATCAACGAAGAATAC 420
    |||||||
DB 383 AGTATCCCATATGATGATAAATTTGATATGCTCCGCCGCTTGCATCAACGAAGAATAC 442
    |||||||

OY 421 CAAAATGATATGTAATTTATGTAATATTCATGATGCGGAATTCGTGCGCAACT 480
    |||||||
DB 443 CAAAATGATATGTAATTTATGTAATATTCATGATGCGGAATTCGTGCGCAACT 502
    |||||||

OY 481 TCGCTGTCTCCAAATTCGATGAGGAGGCTCTCCTAAGCAAGCATCGTTCAAAAGTT 540
    |||||||
DB 503 TCGCTGTCTCCAAATTCGATGAGGAGGCTCTCCTAAGCAAGCATCGTTCAAAAGTT 562
    |||||||

OY 541 GGACAAGCTCTTGAACAT 559
    |||||||
DB 563 GGACAAGCTCTTGAACAT 581
    |||||||

RESULT 3
```

Bj115564 519 bp mRNA linear EST 23-JAN-2002
 LOCUS Bj115564 unpublished oligo-capped cDNA library, C. elegans L1 stage
 DEFINITION Caenorhabditis elegans cDNA clone yk1192h09 5', mRNA sequence.
 ACCESSION Bj115564
 VERSION Bj115564.1 GI:18275658
 KEYWORDS EST.
 SOURCE Caenorhabditis elegans
 ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
 ; Rhabditidae; Peloderinae; Caenorhabditis.
 REFERENCE 1 (bases 1 to 519)
 AUTHORS Kohara, Y., Shin-I, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.
 and Sugano, S.
 A complementary view of the C. elegans genome
 TITLE Unpublished
 JOURNAL Contact: Tadasu Shin-I
 COMMENT Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp.
 FEATURES
 source
 1. 519
 Location/Qualifiers
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /strain="N2"
 /db_xref="taxon:6239"
 /clone="yk1192h09"
 /sex="hermaphrodite"
 /tissue_type="whole animal"
 /dev_stage="L1"
 /clone_lib="unpublished oligo-capped cDNA library, C.
 elegans L1 stage"
 BASE COUNT 170 a 91 c 109 g 149 t
 ORIGIN
 Query Match 30.3%; Score 453; DB 12; Length 519;
 Best local Similarity 99.8%; Pred. No. 1.9e-233;
 Matches 503; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 ATGCTGAATTTCTGAAAAATGTGACAGCTAACAAAAATCGACAGAAAACTCGATTAAG 60
 Db 16 ATGCTGAATTTCTGAAAAATGTGACAGCTAACAAAAATCGACAGAAAACTCGATTAAG 75
 Oy 61 ACCTACTTGTGGGAATCTATTATTCATCATGCTTCGAGAAAGAAAACTTTTCATTCCA 120
 Db 76 ACCTACTTGTGGGAATCTATTATTCATCATGCTTCGAGAAAGAAAACTTTTCATTCCA 135
 Oy 121 GTTGAAGCATTCGAATGTAACCTTACAGTTAATTTTAAGAAATGCGTGAAGGAAGAGTT 180
 Db 136 GTTGAAGCATTCGAATGTAACCTTACAGTTAATTTTAAGAAATGCGTGAAGGAAGAGTT 195
 Oy 181 ATCTTGAAGAAAGTGTGTCATGATTTATGACAAAGAACTGCGATTGATTCAGTCAATGG 240
 Db 196 ATCTTGAAGAAAGTGTGTCATGATTTATGACAAAGAACTGCGATTGATTCAGTCAATGG 255
 Oy 241 TTTCGACGAATGAAAAAGTTTGGCGATTCAGAGTTCTGGCTAGCTTATCGAGCTGAC 300
 Db 256 TTTCGACGAATGAAAAAGTTTGGCGATTCAGAGTTCTGGCTAGCTTATCGAGCTGAC 315
 Oy 301 ACGAAATTTTGGCTCAATATTTTATGCGAGATATGTTGGTTGGCAAAAGCGCGAATG 360
 Db 316 ACGAAATTTTGGCTCAATATTTTATGCGAGATATGTTGGTTGGCAAAAGCGCGAATG 375
 Oy 361 AGTGATCCCAATATGATATAATTTGTATATGCTCCCGCTTCGATTCACGAAGAATAC 420
 Db 376 AGTGATCCCAATATGATATAATTTGTATATGCTCCCGCTTCGATTCACGAAGAATAC 435
 Oy 421 CAAATGATATGTTAAATTTATGTAATTAATGTCATTGATGCGGAAATCTCGCGCCAACT 480
 Db 436 CAAATGATATGTTAAATTTATGTAATTAATGTCATTGATGCGGAAATCTCGCGCCAACT 495

Oy 481 TCCTGTCTCCAAATTCGATGAA 504
 Db 496 TCCTGTCTCCAAATTCGATGAA 519
 RESULT 4
 LOCUS Bj124772 557 bp mRNA linear EST 30-MAY-2003
 DEFINITION Bj124772 unpublished oligo-capped cDNA library, C. elegans L1 stage
 ACCESSION Bj124772
 VERSION Bj124772.2 GI:31246008
 KEYWORDS EST.
 SOURCE Caenorhabditis elegans
 ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
 ; Rhabditidae; Peloderinae; Caenorhabditis.
 REFERENCE 1 (bases 1 to 557)
 AUTHORS Kohara, Y., Shin-I, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.
 and Sugano, S.
 A complementary view of the C. elegans genome
 TITLE Unpublished
 JOURNAL Contact: Tadasu Shin-I
 COMMENT On Jan 23, 2002 this sequence version replaced gi:18284913.
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp.
 FEATURES
 source
 1. 557
 Location/Qualifiers
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /strain="N2"
 /db_xref="taxon:6239"
 /clone="yk1315b11"
 /sex="hermaphrodite"
 /tissue_type="whole animal"
 /dev_stage="L1"
 /clone_lib="unpublished oligo-capped cDNA library, C.
 elegans L1 stage"
 BASE COUNT 180 a 99 c 118 g 158 t 2 others
 ORIGIN
 Query Match 28.9%; Score 433; DB 12; Length 557;
 Best local Similarity 100.0%; Pred. No. 1.3e-222;
 Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 102 AAAAAGCTTCTTCATTCAGTTGACGATTCATTCGTAACCTTACAGTTAATTTAACGA 161
 Db 125 AAAAAGCTTCTTCATTCAGTTGACGATTCATTCGTAACCTTACAGTTAATTTAACGA 184
 Oy 162 ATGCGTGAAGGAAGAGTTATCTTCGAAACAGTGTCCATGATTTATGACAAAGTCCGA 221
 Db 185 ATGCGTGAAGGAAGAGTTATCTTCGAAACAGTGTCCATGATTTATGACAAAGTCCGA 244
 Oy 222 TTTCGATTCAGTCAATGATTTGTCAGCAATTTGAAAAAGTTTGGCGATTCAGAGTTCTGC 281
 Db 245 TTTCGATTCAGTCAATGATTTGTCAGCAATTTGAAAAAGTTTGGCGATTCAGAGTTCTGC 304
 Oy 282 TCGATTTATTCGAGGTGACACGAAATTTTGGCTCAATATTTTATGCGAGCATATGTTGG 341
 Db 305 TCGATTTATTCGAGGTGACACGAAATTTTGGCTCAATATTTTATGCGAGCATATGTTGG 364
 Oy 342 TTTCGCAAAAGCGCGCATGATGATCCCAATATGATTAATAATTTATATGCTCCGCCGT 401
 Db 365 TTTCGCAAAAGCGCGCATGATGATCCCAATATGATTAATAATTTATATGCTCCGCCGT 424
 Oy 402 TGCATTCACGAAGAATATCCAAATATATGTAATTTATGTAATTAATTCATTTGATGCG 461
 Db 425 TGCATTCACGAAGAATATCCAAATATATGTAATTTATGTAATTAATTCATTTGATGCG 484

QY	462	CGAAATGCTGGCCCAACTTCGGCTGCTCCAAAATTCAGTGAAGGGAAGCTCTCCCTAAAG	521
Db	485	CGAATGCTGGCCCAACTTCGGCTGCTCCAAAATTCAGTGAAGGGAAGCTCTCCCTAAAG	544
QY	522	CAAGCATCGTTTC	534
Db	545	CAAGCATCGTTTC	557
RESULT 5			
BJ143186/c			
LOCUS	BJ143186	768 bp	mRNA linear EST 23-JAN-2002
DEFINITION	BJ143186	unpublished oligo-capped cDNA library, C. elegans L1 stage	
ACCESSION	BJ143186	Caenorhabditis elegans cDNA clone yk1192h09 3', mRNA sequence.	
VERSION	BJ143186		
KEYWORDS	EST.		
SOURCE	Caenorhabditis elegans		
ORGANISM	Caenorhabditis elegans		
REFERENCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea		
AUTHORS	Kohara, Y., Shin-I, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.		
TITLE	A complementary view of the C.elegans genome		
JOURNAL	Unpublished		
COMMENT	Contact: Tadashi Shin-I Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6835 Email: tshini@genes.nig.ac.jp.		
FEATURES			
source	1..768	Location/Qualifiers	
	/organism="Caenorhabditis elegans"		
	/mol_type="mRNA"		
	/strain="N2"		
	/db_xref="taxon:6239"		
	/clone="yk1192h09"		
	/sex="hermaphrodite"		
	/tissue_type="whole animal"		
	/dev_stage="L1"		
	/clone_lib="unpublished oligo-capped cDNA library, C. elegans L1 stage"		
BASE COUNT	232 a 151 c 146 g 237 t	2 others	
ORIGIN			
Query Match	27.7%; Score 414; DB 12; Length 768;		
Best Local Similarity	100.0%; Pred. No. 2.9e-212;		
Matches	414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	847	AGATATGACTCAGACGACGTCACATTTGATCAATTTAGCAAAAAGATCCATTTGATCCATG	906
Db	768	AGATATGACTCAGACGACGTCACATTTGATCAATTTAGCAAAAAGATCCATTTGATCCATG	709
QY	907	ATTGAGAGAAAAGTTAAGTTGGAGCAAAAGTTTGGATCATCGAGCCCTTGCTGCTAGCAA	966
Db	708	ATTGAGAGAAAAGTTAAGTTGGAGCAAAAGTTTGGATCATCGAGCCCTTGCTGCTAGCAA	649
QY	967	TTCAATTAACCTCCACGTCGCTTCGATTCATCTCAAAATTTTGCAAAACTGAAGATATCTTAAT	1026
Db	648	TTCAATTAACCTCCACGTCGCTTCGATTCATCTCAAAATTTTGCAAAACTGAAGATATCTTAAT	589
QY	1027	GTGGCAATGATGTCGTCAGATGACACTTAAGACAGTTTTCCTATTCATATCAATTAATGCA	1086
Db	588	GTGGCAATGATGTCGTCAGATGACACTTAAGACAGTTTTCCTATTCATATCAATTAATGCA	529
QY	1087	TTTATGTTCCCGAGTGGTATGCGGAAAAGTATATTTGGAACTGTTCCGCGCAGATGAG	1146
Db	528	TTTATGTTCCCGAGTGGTATGCGGAAAAGTATATTTGGAACTGTTCCGCGCAGATGAG	469

Oy	1147	TTCAAGAACATTCACATGGGATGACATCTGGAGAAAGATCTGCAGAAACCTTACCG	1206
Db	468	TTTCAAGAACATTCACATGGGATGACATCTGGAGAAAGATCTGCAGAAACCTTACCG	409
Oy	1207	CTTACTGTGTTCACAGCCATGCTCTTCCCAAGAGAGATTTAGCAAAATTTAAGGTA	1260
Db	408	CTTACTGTGTTCACAGCCATGCTCTTCCCAAGAGAGATTTAGCAAAATTTAAGGTA	355
RESULT 6			
CB404801/c		607 bp	linear
LOCUS			EST 15-MAY-2003
DEFINITION	OSR028D3_1 AD-wrmcDNA Caenorhabditis elegans CDNA, mRNA sequence.		
ACCESSION	CB404801		
VERSION	CB404801.1	GI:30746528	
KEYWORDS	EST.		
SOURCE	Caenorhabditis elegans		
ORGANISM	Caenorhabditis elegans		
REFERENCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae		
AUTHORS	1 (bases 1 to 607) Reboul,J., Vaglio,P., Rual,J.F., Lamesch,P., Martnez,M., Armstrong C.M., Li,S., Jacotot,L., Bertin,N., Janky,R., Moore,T., Hudson ,J.R., Hatley,J.L., Brasch,M.A., Vandenhaute,J., Boulton,S., Endress,G.A., Jenna,S., Chevet,E., Papasotiriopoulos,V., Tollas,P.P., Placke,J., Snyder,M., Huang,R., Chance,M.R., Lee,H., Doucette-Stamm,L., Hill,D.E. and Vidal,M.		
TITLE	C. elegans ORFome version 1.1: experimental verification of the genome annotation and resource for proteome-scale protein expression		
JOURNAL	Nat. Genet., (2003) In press		
COMMENT	Contact: Vidal M Marc Vidal Laboratory Dana Farber Cancer Institute 1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA Tel: 617 632 5180 Fax: 617 632 5739 Email: Marc.Vidal@dfci.harvard.edu Sequence tag of Gateway entry clones. The primers used were designed on the predicted protein encoding ORF. C. elegans ORFome cloning project : contact david_hill@dfci.harvard.edu or marc_vidal@dfci.harvard.edu POLYA-No.		
FEATURES			
source	location/Qualifiers		
	1..607		
	/organism="Caenorhabditis elegans"		
	/mol_type="mRNA"		
	/strain="N2"		
	/db_xref="taxon:6239"		
	/sex="Hermaphrodite and male"		
	/tissue_type="whole animal"		
	/dev_stage="mixed stage"		
	/clone.lib="AD-wrmcDNA"		
	/note="The AD-wrmcDNA library was generated with poly(A)+ RNA isolated from both hermaphrodite and male N2 worms of all larval stages, embryos, adults and denurs and the subsequent generation of cDNAs by poly(A) priming. The cDNAs were cloned into pPC86"		
BASE COUNT	175 a 131 g 188 t		
ORIGIN			
Query Match	25.7%; Score 385; DB 14; Length 607;		
Best Local Similarity	100.0%; Pred. NO. 1,4e-196;		
Matches	385; Conservative 0; Mismatches 0; Indels 0; Gaps 0		
Oy	876	TCGAATTAGCAAAAGATCCATGATGCCATGATTTGGAGAAAAGTTAAGTTGACAAAA	935
Db	574	TCGAATTAGCAAAAGATCCATGATGCCATGATTTGGAGAAAAGTTAAGTTGACAAAA	515
Oy	936	GTTTAGGTCATCGAGCCCTTGCTGCAGCAATTCATAAACCCTCACGTCGCTTGATTC	995
Db	514	GTTTAGGTCATCGAGCCCTTGCTGCAGCAATTCATAAACCCTCACGTCGCTTGATTC	455

```

QY 996 CAAATTTGCAAAACTGAAGATATCTTATTTGGAGATGATGTCAGATGCATTTGA 1055
    |||
    |||
    |||
Db 454 CAAATTTGCAAAACTGAAGATATCTTATTTGGAGATGATGTCAGATGCATTTGA 395
QY 1056 AGACATTTTCCATTCATATCATATATATACATTTATGTTCCAGTTGTTATGCGAATA 1115
    |||
    |||
    |||
Db 394 AGACATTTTCCATTCATATCATATATATACATTTATGTTCCAGTTGTTATGCGAATA 335
QY 1116 GATATTTTGGAACTGTTCCCGCCAGATGATTCAGAAAGCAATTCAGATGGGATGAATA 1175
    |||
    |||
    |||
Db 334 GATATTTTGGAACTGTTCCCGCCAGATGATTCAGAAAGCAATTCAGATGGGATGAATA 275
QY 1176 CTTGGAGAAAGATATCTGCAGAAACCCCTACCGCTTGACTTTTCAGACCAATGCTTCCCA 1235
    |||
    |||
    |||
Db 274 CTTGGAGAAAGATATCTGCAGAAACCCCTACCGCTTGACTTTTCAGACCAATGCTTCCCA 215
QY 1236 AGAGATTTAGACAAATTTAAGCTA 1260
    |||
    |||
    |||
Db 214 AGAGATTTAGACAAATTTAAGCTA 190

RESULT 7
CA2305
LOCUS C42305 360 bp mRNA linear EST 18-OCT-1999
DEFINITION C42305 Yui1 Kohara unpublished cDNA:Strain N2 hermaphrodite embryo
ACCESSION C42305
KEYWORDS C42305.1 GI:2378542
SOURCE EST.
ORGANISM Caenorhabditis elegans
            Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea
            1 (bases 1 to 360)
REFERENCE Kohara,Y., Motobashi,T., Tabara,H., Watanabe,H., Sugimoto,A., Sano
AUTHORS M., Miyata,A. and Nishigaki,A.
TITLE Expression map of the C.elegans genome
JOURNAL Unpublished
COMMENT Contact: Yui1 Kohara
            Genome Biology Lab.
            National Institute of Genetics
            Yata 1111, Mishima, Shizuoka 411, Japan
            Tel: 81-559-81-6854
            Fax: 81-559-81-6855
            Email: ykohara@lab.nig.ac.jp.
            Location/Qualifiers
FEATURES
    source
        1..360
        /organism="Caenorhabditis elegans"
        /mol_type="mRNA"
        /strain="N2"
        /db_xref="taxon:6239"
        /clone="yk292d9"
        /sex="hermaphrodite"
        /dev_stage="embryo"
        /clone_lib="Yui1 Kohara unpublished cDNA:Strain N2
            hermaphrodite embryo"
BASE COUNT 119 a 62 c 78 g 101 t
ORIGIN
Query Match 24.0%; Score 360; DB 14; Length 360;
Best Local Similarity 100.0%; Pred. No. 4e-183;
Matches 360; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Db 121 TACAGTTAAATTTTACGAATGCGTGAAGAGGATTAATCTTCGAAACAGTGGCCATGA 180
QY 204 TTATGACAAAGACTGGATTCGATTCAGTCAGATGTTGGCAAGTAATGAAGAACTTGG 263
    |||
    |||
    |||
Db 181 TTATGACAAAGACTGGATTCGATTCAGTCAGATGTTGGCAAGTAATGAAGAACTTGG 240
QY 264 CGGATCAGAGTTCTGCTCAGTTTATCGAGCTGACAGAAATTTGGCTCAATATTTT 323
    |||
    |||
    |||
Db 241 CGGATCAGAGTTCTGCTCAGTTTATCGAGCTGACAGAAATTTGGCTCAATATTTT 300
QY 324 ATCGACGATATGTTGGTTTGGCAAAAGCCGCAATGATGATCCCAATATGATTAAT 383
    |||
    |||
    |||
Db 301 ATCGACGATATGTTGGTTTGGCAAAAGCCGCAATGATGATCCCAATATGATTAAT 360

RESULT 8
BUI29361/c
LOCUS BUI29361 761 bp mRNA linear EST 23-JAN-2002
DEFINITION BUI29361 unpublished oligo-capped cDNA library, C. elegans L1 stage
ACCESSION BUI29361
KEYWORDS BUI29361.1 GI:18289518
SOURCE EST.
ORGANISM Caenorhabditis elegans
            Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea
            1 (bases 1 to 761)
REFERENCE Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
AUTHORS and Sugano,S.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished
COMMENT Contact: Tadasi Shin-I
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshin@genes.nig.ac.jp.
            Location/Qualifiers
FEATURES
    source
        1..761
        /organism="Caenorhabditis elegans"
        /mol_type="mRNA"
        /strain="N2"
        /db_xref="taxon:6239"
        /clone="yk1029h03"
        /sex="hermaphrodite"
        /tissue.type="whole animal"
        /dev_stage="L1"
        /clone_lib="unpublished oligo-capped cDNA library, C.
            elegans L1 stage"
BASE COUNT 233 a 150 c 144 g 232 t 2 others
ORIGIN
Query Match 23.8%; Score 356; DB 12; Length 761;
Best Local Similarity 99.8%; Pred. No. 7.6e-181;
Matches 406; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY		1094	TCCAGTTGGTATTGCGGAAGAATATAATTGGAACTGTTCGCCGAGTAGTTCAAAG	1153
Db		521	TCCAGTTGGTATTGCGGAAGAATATAATTGGAACTGTTCGCCGAGTAGTTCAAAG	462
OY		1154	GAACATTCAATGGATGTAATTAATTGGAGAAGAATTCGAGAAAACCTCAGCTTGACT	1213
Db		461	GAACATTCAATGGATGTAATTAATTGGAGAAGAATTCGAGAAAACCTCAGCTTGACT	402
OY		1214	TGTTCAAGCCCAATGCCCTTCCCAGAGAGATTAGCAATTTAAGCTA	1260
Db		401	TGTTCAAGCCCAATGCCCTTCCCAGAGAGATTAGCAATTTAAGCTA	355
RESULT 9	AU216908/c			
LOCUS	AU216908	677 bp	mRNA	Linear EST 17-JUL-2001
DEFINITION	AU216908 unpublished oligo-capped cDNA library, stage L4			
ACCESSION	Caeenorhabditis elegans cDNA clone YK843e05 3', mRNA sequence.			
VERSION	AU216908			
KEYWORDS	AU216908.1 GI:14855065			
SOURCE	EST.			
ORGANISM	Caeenorhabditis elegans			
REFERENCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea			
AUTHORS	1 (bases 1 to 677) Kohara,Y., Shih-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,I.Y. and Sugano,S.			
TITLE	A complementary view of the C.elegans genome			
JOURNAL	Unpublished			
COMMENT	Contact: Yuji Kohara Genome Biology Lab. National Institute of Genetics Yata 1111, Mishima, Shizuoka 411, Japan Tel: 81-559-81-6854 Fax: 81-559-81-6855 Email: ykoha@elab.nig.ac.jp.			
FEATURES	Location/Qualifiers			
source	1..677			
	/organism="Caenorhabditis elegans"			
	/mol_type="mRNA"			
	/strain="N2"			
	/db_xref="taxon:6239"			
	/clone="YK843e05"			
	/sex="Hermaphrodite"			
	/tissue_type="whole animal"			
	/dev_stage="L4"			
	/clone_lib="unpublished oligo-capped cDNA library, stage L4"			
BASE COUNT	210 a 136 c 130 g 201 t			
ORIGIN				
	Query Match 22.3%; Score 334; DB 9; Length 677;			
	Best Local Similarity 100.0%; Pred. No. 6e-169;			
	Matches 334; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
OY		927	TGCACAAAAGTTTGAAGTCATCGACCCTTGCTGACGAATTCATTAACCTCCAGCTGCC	986
Db		677	TGCACAAAAGTTTGAAGTCATCGACCCTTGCTGACGAATTCATTAACCTCCAGCTGCC	618
OY		987	TTGCATTCTCAAAATTTTGCACAAACCTGAGAGATATCTTATTTGTGGGAATGATGGTGCAGA	1046
Db		617	TTGCATTCTCAAAATTTTGCACAAACCTGAGAGATATCTTATTTGTGGGAATGATGGTGCAGA	558
OY		1047	TGCACCTGAGACAGTTTTTCCATTATCATCAATTAATCAATTATGTGCCAGTTGGTTA	1106
Db		557	TGCACCTGAGACAGTTTTTCCATTATCATCAATTAATCAATTATGTGCCAGTTGGTTA	498
OY		1107	TGCGGAAAAGTATTAATTGGAACTGTTCCGCCAGATGAGTTCAAAGACAATTCAGATG	1166
Db		497	TGCGGAAAAGTATTAATTGGAACTGTTCCGCCAGATGAGTTCAAAGACAATTCAGATG	438

QY	1167	GGATGATCTACTGTGGAGAAAGAAATCTGGAGAAACCCTACCGCTTACCTGTTCAGACCAT	1226
Db	437	GGATGAATCTACTGGAGAAAGAAATCTGGAGAAACCCTACCGCTTACCTGTTCAGACCAT	378
QY	1227	GCCTTCCCAAGAGAGATTTAGCAAAATTTAAAGTA	1260
Db	377	GCCTTCCCAAGAGAGATTTAGCAAAATTTAAAGTA	344
RESULT 10			
LOCUS	AV197399	380 bp	linear
DEFINITION	AV197399 Yuji Kohara unpublished cDNA:Strain N2 hermaphrodite		EST 26-JUL-1999
ACCESSION	AV197399		
VERSION	AV197399.1		
KEYWORDS	embryo Caenorhabditis elegans cDNA clone yk65h11 5', mRNA		
SOURCE	sequence.		
ORGANISM	Caenorhabditis elegans		
REFERENCE	Caenorhabditis elegans		
AUTHORS	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea; Rhabditidae; Peloderae; Caenorhabditis.		
	1 (bases 1 to 380)		
	Kohara, Y., Shin-I, T., Thierry-Mieg, D., Thierry-Mieg, D., Mitsuki, H., Sano, Nishigaki, A., Motobashi, T., Zeng, Q., Matanabe, H., Sugimoto, A., Sano, M., Miyata, A., Mitani, Y., Iida, K., Uesugi, H., Sugiyama, Y. and Nomoto, H.		
TITLE	Expressed genes in C.elegans		
JOURNAL	Unpublished		
COMMENT	Contact: Yuji Kohara Genome Biology Lab. National Institute of Genetics Yata 1111, Mishima, Shizuoka 411, Japan Tel: 81-559-81-6854 Fax: 81-559-81-6855 Email: ykohara@lab.nig.ac.jp.		
FEATURES	Location/Qualifiers		
SOURCE	1..380		
	/organism="Caenorhabditis elegans"		
	/mol_type="mRNA"		
	/strain="N2"		
	/db_xref="taxon:6239"		
	/clone="yk65h11"		
	/sex="hermaphrodite"		
	/dev_stage="embryo"		
	/clone_lib="Yuji Kohara unpublished cDNA:Strain N2 hermaphrodite embryo"		
BASE COUNT	125 a 63 c 83 g 108 t 1 others		
ORIGIN			
Query Match	21.5%; Score 322; DB 9; Length 380;		
Best Local Similarity	100.0%; Pred. No. 1.0e-152;		
Matches 322;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	71	GGGATCTCTATTTCACATCAGTTCCAGAAAGAAAACTTTCTTCATCCAGTTGAAGCAT	130
Db	59	GGGATCTCTATTTCACATCAGTTCCAGAAAGAAAACTTTCTTCATCCAGTTGAAGCAT	118
QY	131	TCAATCGTAACCTTACAGAGTTAATTTTAAACGAATGCGTGAAGGAAGAGTTATCTTCGAAA	190
Db	119	TCAATCGTAACCTTACAGAGTTAATTTTAAACGAATGCGTGAAGGAAGAGTTATCTTCGAAA	178
QY	191	CAGTGGTCCATGATTTAGCAAGAAGCTCGATTGCAATTCAGTCAGATGCTTTCACAGAA	250
Db	179	CAGTGGTCCATGATTTAGCAAGAAGCTCGATTGCAATTCAGTCAGATGCTTTCACAGAA	238
QY	251	TTGAAAAAGTTGGCGGATACAGAGTTCTGCGCTCACTTATCGGAGCTGACACGAATTTT	310
Db	239	TTGAAAAAGTTGGCGGATACAGAGTTCTGCGCTCACTTATCGGAGCTGACACGAATTTT	298
QY	311	GGGTCATATTTTATTCGAGATATATGTTGGTTTGGCAACGCCCAATGATGATCCCA	370
Db	299	GGGTCATATTTTATTCGAGATATATGTTGGTTTGGCAACGCCCAATGATGATCCCA	358

OY 371 ATATGATAAATGTATATGC 392
 DB 359 ATATGATAAATGTATATGC 380

RESULT 11

LOCUS AU111588 340 bp mRNA linear EST 19-OCT-2000
 DEFINITION AU111588 unpublished oligo-capped cDNA library Caenorhabditis
 elegans cDNA clone yk732e5 5', mRNA sequence.

ACCESSION AU111588
 VERSION AU111588.1 GI:10925155
 KEYWORDS EST.
 SOURCE Caenorhabditis elegans
 ORGANISM Caenorhabditis elegans

REFERENCE Eukaryota; Metazoa; Nematoda; Chromodorea; Rhabditida; Rhabditoidea
 ; Rhabditidae; Pelodierinae; Caenorhabditis.
 1 (bases 1 to 340)
 Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
 and Sugano,S.

TITLE A complementary view of the C. elegans genome
 JOURNAL Unpublished
 COMMENT Contact: Yuji Kohara
 Genome Biology Lab.
 National Institute of Genetics
 Yata 1111, Mishima, Shizuoka 411, Japan
 Tel: 81-559-81-6854
 Fax: 81-559-81-6855
 Email: ykohara@lab.nig.ac.jp.

FEATURES
 source
 1..340
 Location/Qualifiers
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /strain="N2"
 /db_xref="taxon:6239"
 /clone="yk732e5"
 /sex="Hermaphrodite"
 /tissue_type="whole animal"
 /dev_stage="varied"
 /clone_lib="unpublished oligo-capped cDNA library"

BASE COUNT 111 a 58 c 72 g 98 t 1 others
 ORIGIN

Query Match 20.6%; Score 308; DB 9; Length 340;
 Best Local Similarity 100.0%; Pred. No. 5,7e-155;
 Matches 308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 GTCTGAATTTCTGAAATTTGTCAGAGCTAACAAAAAATCGACAGAAAACTCGATAGAC 62
 DB 33 GTCTGAATTTCTGAAATTTGTCAGAGCTAACAAAAAATCGACAGAAAACTCGATAGAC 92
 OY 63 CTACTTGGGGAATCTATTACATCAGTCGAGAAAGAAACTCTTCATCCAGT 122
 DB 93 CTACTTGGGGAATCTATTACATCAGTCGAGAAAGAAACTCTTCATCCAGT 152
 OY 123 TGAACATTCATCTTAACCTTACAGTAAATTTAAACGAATCGTGAAGAGAGATTAT 182
 DB 153 TGAACATTCATCTTAACCTTACAGTAAATTTAAACGAATCGTGAAGAGAGATTAT 212
 OY 183 CTTCGAAACAGTGTCCATGATTATGACAGAACTGCGATTCGATTCAAGTCAGATGTT 242
 DB 213 CTTCGAAACAGTGTCCATGATTATGACAGAACTGCGATTCGATTCAAGTCAGATGTT 272
 OY 243 TGCAGCAATTTGAAAGTTTGGGATACAGAGTTGCTGCTGTTATTCGGAGTCGACAC 302
 DB 273 TGCAGCAATTTGAAAGTTTGGGATACAGAGTTGCTGCTGTTATTCGGAGTCGACAC 332
 OY 303 GAAATTTT 310
 DB 333 GAAATTTT 340

RESULT 12
 LOCUS BJ153087/c 686 bp mRNA linear EST 24-JAN-2002
 DEFINITION BJ153087 unpublished oligo-capped cDNA library, C. elegans L1 stage
 Caenorhabditis elegans cDNA clone yk1315b11 3', mRNA sequence.

ACCESSION BJ153087
 VERSION BJ153087.1 GI:18321072
 KEYWORDS EST.
 SOURCE Caenorhabditis elegans
 ORGANISM Caenorhabditis elegans

REFERENCE Eukaryota; Metazoa; Nematoda; Chromodorea; Rhabditida; Rhabditoidea
 ; Rhabditidae; Pelodierinae; Caenorhabditis.
 1 (bases 1 to 686)
 Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
 and Sugano,S.

TITLE A complementary view of the C. elegans genome
 JOURNAL Unpublished
 COMMENT Contact: Yadao Shin-I
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp.

FEATURES
 source
 1..686
 Location/Qualifiers
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /strain="N2"
 /db_xref="taxon:6239"
 /clone="yk1315b11"
 /sex="hermaphrodite"
 /tissue_type="whole animal"
 /dev_stage="L1"
 /clone_lib="unpublished oligo-capped cDNA library, C.
 elegans L1 stage"

BASE COUNT 214 a 133 c 131 g 204 t 4 others
 ORIGIN

Query Match 20.4%; Score 306; DB 12; Length 686;
 Best Local Similarity 100.0%; Pred. No. 8.9e-154;
 Matches 306; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 955 TTGGCTCAGCAATTCATTAACCTCCAGCGCTGCTGATCTCAATTTGGCAAACTGAA 1014
 DB 666 TTGGCTCAGCAATTCATTAACCTCCAGCGCTGCTGATCTCAATTTGGCAAACTGAA 607
 OY 1015 GGATATCTTATTTGGGGAATGAGTGTCCAGATGCATTTGAAGACAGTTTCTATTCAT 1074
 DB 606 GGATATCTTATTTGGGGAATGAGTGTCCAGATGCATTTGAAGACAGTTTCTATTCAT 547
 OY 1075 ATCAATTAATACATTTATGTTCCAGTGTGTTATGCGGAAAGTATTAATTTGGACCTGTT 1134
 DB 546 ATCAATTAATACATTTATGTTCCAGTGTGTTATGCGGAAAGTATTAATTTGGACCTGTT 487
 OY 1135 CCGCCAGATGAGTTCAAGGAACATTCAGATGAGTGAATACTTGAGAAAGATCTGCA 1194
 DB 486 CCGCCAGATGAGTTCAAGGAACATTCAGATGAGTGAATACTTGAGAAAGATCTGCA 427
 OY 1195 GAAACCTTACCGCTTGACTGTTCGAAGCCATGCTTCCCAAGAGAGATTAGCAAAATTT 1254
 DB 426 GAAACCTTACCGCTTGACTGTTCGAAGCCATGCTTCCCAAGAGAGATTAGCAAAATTT 367
 OY 1255 AAGGTA 1260
 DB 366 AAGGTA 361

RESULT 13
 LOCUS AU116296/c 300 bp mRNA linear EST 19-OCT-2000
 DEFINITION AU116296 unpublished oligo-capped cDNA library Caenorhabditis
 elegans cDNA clone yk743a4 3', mRNA sequence.

ACCESSION	A0116296		
VERSION	A0116296.1		
KEYWORDS	GI:10929863		
SOURCE	EST.		
ORGANISM	Caenorhabditis elegans		
REFERENCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea		
AUTHORS	1 (bases 1 to 300) Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y. and Sugano,S.		
TITLE	A complementary view of the C. elegans genome		
JOURNAL	Unpublished		
COMMENT	Contact: Yuji Kohara Genome Biology Lab. National Institute of Genetics Yata 1111, Mishima, Shizuoka 411, Japan Tel: 81-559-81-6854 Fax: 81-559-81-6855 Email: ykohara@lab.nig.ac.jp. location/Qualifiers		
FEATURES	1..300		
source	/organism="Caenorhabditis elegans" /mol_type="mRNA" /strain="N2" /db_xref="taxon:6239" /clone="yk743a4" /sex="Hermaphrodite" /tissue_type="whole animal" /dev_stage="varied" /clone_lib="unpublished oligo-capped cDNA library"		
BASE COUNT	95 a 63 c 34 g 88 t		
ORIGIN			
Query Match	14.7%; Score 220; DB 9; Length 300;		
Best Local Similarity	100.0%; Pred. No. 2.5e-107;		
Matches	220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
OY	1278 GGTGAGACTACGCGCTTGAAGCTGCTGACATGTGTGAAATAAGTTATTGTCACAGTAC 1337		
Db	300 GGTGAGACTACGCGCTTGAAGCTGCTGACATGTGTGAAATAAGTTATTGTCACAGTAC 241		
OY	1338 AGTAATAATCATGTCATGGAAGACATGATAAATGTCAATTTGCGACGGCTGGATGAAGAA 1397		
Db	240 AGTAATAATCATGTCATGGAAGACATGATAAATGTCAATTTGCGACGGCTGGATGAAGAA 181		
OY	1398 TGAATGAACGTATGATGTGACATCCCATGATATTCTACCCGATAGAGTGTGAAGCGCA 1457		
Db	180 TGAATGAACGTATGATGTGACATCCCATGATATTCTACCCGATAGAGTGTGAAGCGCA 121		
OY	1458 CAGTTATGTTTACACCTCCGAAAAGTCAACTATTGA 1497		
Db	120 CAGTTATGTTTACACCTCCGAAAAGTCAACTATTGA 81		
RESULT 14			
A0115428/c	300 bp mRNA linear EST 19-OCT-2000		
LOCUS	A0115428 unpublished oligo-capped cDNA library Caenorhabditis		
DEFINITION	elegans cDNA clone yk7325 3', mRNA sequence.		
ACCESSION	A0115428		
VERSION	A0115428.1		
KEYWORDS	EST.		
SOURCE	Caenorhabditis elegans		
ORGANISM	Caenorhabditis elegans		
REFERENCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea		
AUTHORS	1 (bases 1 to 300) Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y. and Sugano,S.		
TITLE	A complementary view of the C. elegans genome		
JOURNAL	Unpublished		
COMMENT	Contact: Yuji Kohara Genome Biology Lab.		

FEATURES		source		National Institute of Genetics Yata 1111, Mishima, Shizuoka 411, Japan Tel: 81-559-81-6854 Fax: 81-559-81-6855 Email: ykohara@elab.nig.ac.jp. Location/Qualifiers	
FEATURES		source		1. 300 /organism="Caenorhabditis elegans" /mol_type="mRNA" /strain="N2" /db_xref="taxon:6239" /clone="YK7325" /sex="Hermaphrodite" /tissue_type="whole animal" /dev_stage="varied" /clone_idb="unpublished oligo-capped cdna library"	
BASE COUNT		94 a 60 c 54 g 89 t 3 others			
ORIGIN					
Query Match		14.4% Score 215; DB 9; Length 300;			
Best Local Similarity		100.0%; Pred. No.1.3e-104;			
Matches 215; Conservative		0; Mismatches 0; Indels 0; Gaps 0;			
OY	1283	GACTAGCCCTTGAAGCTGCTGCACATGTGTGAAAAATCAGTTATTATTTGCCAGCTACAGTGA	1342		
Db	300	GACTAGCCCTTGAAGCTGCTGCACATGTGTGAAAAATCAGTTATTATTTGCCAGCTACAGTGA	241		
OY	1343	AATCAGTTTCATGGAAGACGTATTAATGTCATATTCGACGGCTGGAGTGAAGAATTGTATG	1402		
Db	240	AATCAGTTTCATGGAAGACGTATTAATGTCATATTCGACGGCTGGAGTGAAGAATTGTATG	181		
OY	1403	AACCTGATGATGTGCACTCCCATGATATTTCTACCGATGAGATGCTGTGAAGCCGACAGTT	1462		
Db	180	AACCTGATGATGTGCACTCCCATGATATTTCTACCGATGAGATGCTGTGAAGCCGACAGTT	121		
OY	1463	ATGTTTACACCTCCGAAAAAGTACACATTTGA	1497		
Db	120	ATGTTTACACCTCCGAAAAAGTACACATTTGA	86		
RESULT 15		337 bp mRNA linear EST 30-MAY-2003			
A0112450					
LOCUS					
DEFINITION		A0112450 unpublished oligo-capped cdna library Caenorhabditis			
ACCESSION		A0112450			
VERSION		A0112450.2 GI:31237498			
KEYWORDS		EST.			
SOURCE		Caenorhabditis elegans			
ORGANISM		Caenorhabditis elegans			
REFERENCE		Eukaryotes; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabdicoidea			
AUTHORS		; Rhabdilitidae; Pelodierinae; Caenorhabditis.			
TITLE		1 (bases 1 to 337)			
JOURNAL		Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.			
COMMENT		A complementary view of the C. elegans genome			
		Unpublished			
		On Oct 19, 2000 this sequence version replaced gi:10926017.			
		Contact: Yuji Kohara			
		Genome Biology Lab.			
		National Institute of Genetics			
		Yata 1111, Mishima, Shizuoka 411, Japan			
		Tel: 81-559-81-6854			
		Fax: 81-559-81-6855			
		Email: ykohara@elab.nig.ac.jp.			
		Location/Qualifiers			
FEATURES		1. 337			
source		/organism="Caenorhabditis elegans"			
		/mol_type="mRNA"			
		/strain="N2"			
		/db_xref="taxon:6239"			
		/clone="YK7434"			
		/sex="Hermaphrodite"			

/tissue_type="whole animal"
/dev_stage="varied"
/clone_lib="unpublished oligo-capped cDNA library"
BASE COUNT 107 a 59 c 74 g 92 t 5 others
ORIGIN

Query Match 12.4%; Score 186; DB 9; Length 337;
Best Local Similarity 100.0%; Pred. No. 6.9e-89;
Matches 186; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCGAATTTCTGAAATTTGTCAGAGCTAACAAAAATCGACAGAAACTCGATAAG 60
|||
Db 23 ATGTCGAATTTCTGAAATTTGTCAGAGCTAACAAAAATCGACAGAAACTCGATAAG 82
|||
QY 61 ACCTACTGTGGGAATCCATTTACATCACTTCGAGAAAGGAAACTTCTTCATTCCA 120
|||
Db 83 ACCTACTGTGGGAATCCATTTACATCACTTCGAGAAAGGAAACTTCTTCATTCCA 142
|||
QY 121 GTTGAAGCATTCATCGTACCTTACAGTTAATTTTAAACGAATCGTGAAGAGAGATT 180
|||
Db 143 GTTGAAGCATTCATCGTACCTTACAGTTAATTTTAAACGAATCGTGAAGAGAGATT 202
|||
QY 181 ATCTTC 186
|||
Db 203 ATCTTC 208

Search completed: September 2, 2003, 20:31:03
Job time : 3446 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2003, 16:44:20 ; Search time 5682 Seconds

(without alignments) 10778.196 Million cell updates/sec

Title: US-09-872-523-6

Perfect score: 1497
Sequence: 1 atgctgattctcgtgaaat.....cgaaaagcacactatgta 1497

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 2045481386 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Database :

GenBankl:*

1: gb_da:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_da:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rtd:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query			DB ID	Description
		Match	Length	DB		
C 1	833	55.6	30911	3	CER06C7	271266 Caenorhabd
C 2	25	1.7	218764	2	AC103229	AC103229 Rattus no
C 3	23	1.5	133343	2	AP005398	AP005398 Oryza sat
C 4	22	1.5	157654	2	AC091791	AC091791 Sus scrof
C 5	22	1.5	183239	2	AC128697	AC128697 Mus muscu
C 6	22	1.5	193311	2	AC118694	AC118694 Mus muscu
C 7	22	1.5	195037	2	AC091790	AC091790 Sus scrof
C 8	22	1.5	205682	2	AC140183	AC140183 Mus muscu
C 9	22	1.5	229331	2	AC118916	AC118916 Rattus no
C 10	22	1.5	231442	2	AC094981	AC094981 Rattus no
C 11	22	1.5	253661	2	AC128235	AC128235 Rattus no
C 12	21	1.4	872	6	BD146582	BD146582 Primer fo
C 13	21	1.4	2202	9	BD159393	BD159393 Primer fo
C 14	21	1.4	2202	9	AK022163	AK022163 Homo sapi
C 15	21	1.4	11822	3	DMC137E7	AL021108 Drosophill
C 16	21	1.4	42052	3	U80029	U80029 Caenorhabd
C 17	21	1.4	59824	2	AC109340	AC109340 Homo sapi
C 18	21	1.4	110000	2	AC105643	Continuation (5 of
C 19	21	1.4	135580	2	AC017687	AC017687 Drosophill
C 20	21	1.4	140658	5	AF112374	AF112374 Dantio rer
C 21	21	1.4	146074	2	AC141806	AC141806 Apis mell
C 22	21	1.4	152575	2	AC141690	AC141690 Apis mell
C 23	21	1.4	166231	2	AC120880	AC120880 Mus muscu
C 24	21	1.4	170490	3	AC104147	AC104147 Drosophill
C 25	21	1.4	173525	2	AC067893	AC067893 Homo sapi
C 26	21	1.4	175262	3	AC098576	AC098576 Drosophill
C 27	21	1.4	176709	9	AC011737	AC011737 Homo sapi
C 28	21	1.4	184585	9	AC079465	AC079465 Homo sapi
C 29	21	1.4	184649	2	AC025185	AC025185 Homo sapi
C 30	21	1.4	187003	2	AC023353	AC023353 Homo sapi
C 31	21	1.4	196745	9	AC087431	AC087431 Homo sapi
C 32	21	1.4	204478	2	AC130948	AC130948 Rattus no
C 33	21	1.4	205066	2	AC120803	AC120803 Rattus no
C 34	21	1.4	220622	2	AC118570	AC118570 Lemur cat
C 35	21	1.4	240918	2	AC137254	AC137254 Rattus no
C 36	21	1.4	243833	2	AC097386	AC097386 Rattus no
C 37	21	1.4	251237	2	AC105634	AC105634 Rattus no
C 38	21	1.4	261251	2	AC096413	AC096413 Rattus no
C 39	21	1.4	261603	2	AC098318	AC098318 Rattus no
C 40	21	1.4	300933	3	AE003422	AE003422 Drosophill
C 41	21	1.4	323991	2	AC098512	AC098512 Rattus no
C 42	20	1.3	2044	6	AR302542	AR302542 Sequence
C 43	20	1.3	2863	6	AR302510	AR302510 Sequence
C 44	20	1.3	2863	6	AX058963	AX058963 Sequence
C 45	20	1.3	3646	6	AR302540	AR302540 Sequence

ALIGNMENTS

RESULT 1
CER06C7 30911 bp DNA linear INV 21-MAY-2003
LOCUS CER06C7/c
DEFINITION Caenorhabditis elegans cosmid R06C7, complete sequence.
ACCESSION Z71266
VERSION Z71266.1 GI:1279324
KEYWORDS HTG; Adenylosuccinate lyase; Myosin heavy chain; Rat TOAD-64
SOURCE Protein like; Serine/threonine-protein kinase; Zinc finger protein.
ORGANISM Caenorhabditis elegans
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderiinae; Caenorhabditis.
REFERENCE none.
AUTHORS Genome sequence of the nematode C. elegans: a platform for

Pred. No. is the number of results predicted by chance to have a

JOURNAL Investigating biology. The C. elegans Sequencing Consortium
MEDLINE Science 282 (5396), 2012-2018 (1998)
99069613
REMARK The C.elegans Sequencing Consortium.
AUTHORS 2 (bases 1 to 30911)
TITLE Gardner,A.E.
JOURNAL Direct Submission

COMMENT Submitted (19-APR-1996) Nematode Sequencing Project, Sanger
Institute, Hinxton, Cambridge CB10 1SA, England and Department of
Genetics, Washington University, St. Louis, MO 63110, USA. E-mail:
j.sanger@ac.uk or r.w.nematode@wustl.edu
Coding sequences below are predicted from computer analysis, using
predictions from Genefinder (P. Green, U. Washington), and other
available information.

Current sequence finishing criteria for the C. elegans genome
sequencing consortium are that all bases are either sequenced
unambiguously on both strands, or on a single strand with both a
dye primer and dye terminator reaction, from distinct subclones.
Exceptions are indicated by an explicit note.

IMPORTANT: This sequence is not the entire insert of clone R06C7.
It may be shorter because we only sequence overlapping sections
once, or longer because we arrange for a small overlap between
neighbouring submissions.

The true left end of clone R06C7 is at 1 in this sequence. The true
right end of clone R06C7 is at 8718 in
sequence Z71261.

The true left end of clone F21C3 is at 30808 in this sequence. The
start of this sequence (1..106) overlaps with the end of sequence
Z98261.

The end of this sequence (30808..30911) overlaps with the start of
sequence Z71261.

For a graphical representation of this sequence and its analysis
see: - [http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?](http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?name=R06C7)
name=R06C7

IMPORTANT: This sequence is NOT necessarily the entire insert of
the specified clone. It may be shorter because we only sequence
overlapping sections once, or longer because we arrange for a small
overlap between neighbouring submissions.

FEATURES

source

1..30911

/organism="Caenorhabditis elegans"

/mol_type="genomic DNA"

/strain="Bristol N2"

/db_xref="taxon:6239"

/chromosome="1"

/clone="R06C7"

join(complement(588..775),complement(395..539),
complement(169..295),complement(298261.1:204..368),
complement(298261.1:105..155),
complement(271258.1:40294..40359),
complement(271258.1:40101..40244),
complement(271258.1:39857..39994),
complement(271258.1:39258..39701),
complement(271258.1:38262..39212),
complement(271258.1:38105..38207),
complement(271258.1:37940..38057),
complement(271258.1:37773..37895))

/gene="C01H6.9"

join(complement(588..775),complement(395..539),
complement(169..295),complement(298261.1:204..368),
complement(298261.1:105..155),
complement(271258.1:40294..40359),
complement(271258.1:40101..40244),
complement(271258.1:39857..39994),
complement(271258.1:39258..39701),
complement(271258.1:38262..39212),
complement(271258.1:38105..38207),
complement(271258.1:37940..38057),
complement(271258.1:37773..37895))

CDS

join(complement(588..775),complement(395..539),
complement(169..295),complement(298261.1:204..368),
complement(298261.1:105..155),
complement(271258.1:40294..40359),
complement(271258.1:40101..40244),
complement(271258.1:39857..39994),
complement(271258.1:39258..39701),
complement(271258.1:38262..39212),
complement(271258.1:38105..38207),
complement(271258.1:37940..38057),
complement(271258.1:37773..37895))
/gene="C01H6.9"
/standard_name="C01H6.9"
/note="contains similarity to Pfam domain: PF00069

(Eukaryotic protein kinase domain), Score=37.5,

E-value=8.2e-10, N=2

cdna EST yk42h5.3 comes from this gene

cdna EST yk42h5.5 comes from this gene"

/codon_start=1

/product="Hypothetical protein C01H6.9"

/protein_id="CAA5847.1"

/db_xref="GI:3878842"

/db_xref="GOA:017985"

/db_xref="SPTRMBL:017985"

/translation="MPKPRIKVPISGKKARNFARNRNLSSVITTECPIDPEPE

FSKPALFSPNDPERRPKKKPKPPVPPSPDVSDESDPAAOGKPDADA

DEBKIKIKIEAPSLDVSLEPVNDGIRKIKENGHIITHDROPZAMPREPAKTD

FAKMEOROKPVKQVPEMNSKANQVCPINVAFCGSIILADHRKKSLAPSLF

RVEGTPORSKMIIPAKHVOCODERLAIYSTPIPVNTAODPTGSIYFSPQNA

STPIEKOTAGRPKSPFPVPLVNDSCENTLKSCFEOJAOXSANDQKETTINNPIE

EYSRGVHSSSLAGSIQOPODLEETALISENLTSTKTPTIESRNSHNSKIMTQO

VSDVEDVEDYTFNFKRFTSVTELLHEDEHRCQDLNLSKSKYKKRQPKSPETMO

GMSYVMQDDDDLEALFEIEKNEIKRTTLOPOOROPSSRSDINSAMEBSLQO

FLEDTWEEFGKTYSESRAESRNIPTGMTIHNDPSILPEYLEDLSEIDPSMOL

LHVGOEKSTWISLPKSLADGRVKRIGGAYGEVSTIWDGPAVAKIYPEKDCG

NROFGYHSEEMQTSVDVLEPVYVMKELALRDEDMNSTPNFIEISAIVWAKP

KGLSAMDSDYDKLESENTPRDVYSIDONITLFSANGGIALDEFLSESENELFTII

HOLVLSMNAEALAEFPHRDLHGNTLIDNGVYELVYTHGOKVPLSTGIKVNIID

FTLSRISKATVYWDENDPAIFEGODDPOFVYRMRNRCKRKKRRTLMWT

VYIANRLIDTRICKGLITERKRMELVLDRREFSCGSLTNEFFSDYEGPIG

MSWTRQ"

complement(join(1315..1480,1528..1715,1747..1783,

1834..2156))

/gene="R06C7.2"

complement(join(1315..1480,1528..1715,1747..1783,

1834..2156))

/gene="R06C7.2"

/standard_name="R06C7.2"

/codon_start=1

/product="Hypothetical protein R06C7.2"

/protein_id="CAA5841.1"

/db_xref="GI:3878836"

/db_xref="SPTRMBL:021772"

/translation="WFSALVHDSRRDCKKRYQSOTSHONYFLKYSENLGRRKING

NFLHLFHAERKVLFOYVLMKDQTKTKRQETQCLSKLANEENNRKRIKODERN

KKLRLINSLKLVSSPENLPQILFELISENLELERRHFCAFTFKLSDPOLF

FPMPYVANYEAKSGRNLWNEAIIADAIPEVSGEARMQOITDSIKKSGFLIDBR

VMSALIIQNSCSIFNSNF"

join(3063..3227,3274..3854,3899..4631,4680..4855,

4899..5074,5117..5463,5510..5749,5790..6097,6143..6254)

/gene="R06C7.1"

join(3063..3227,3274..3854,3899..4631,4680..4855,

4899..5074,5117..5463,5510..5749,5790..6097,6143..6254)

/gene="R06C7.1"

/standard_name="R06C7.1"

/note="contains similarity to Pfam domain: PF02170 (ZAP

domain), Score=151.7, E-value=4.2e-42, N=1; PF02171 (Plw1

domain), Score=456.7, E-value=6.3e-134, N=1

cdna EST yk31a12.5 comes from this gene

cdna EST yk21g1.3 comes from this gene

cdna EST yk21g1.3 comes from this gene

cdna EST yk31a12.3 comes from this gene

cdna EST yk36g4.3 comes from this gene

cdna EST yk31a4.3 comes from this gene

cdna EST yk31a4.5 comes from this gene

cdna EST yk93h11.5 comes from this gene

cdna EST yk125b7.5 comes from this gene

cdna EST yk125b7.3 comes from this gene

cdna EST yk236e4.3 comes from this gene

cdna EST yk295d11.3 comes from this gene

cdna EST yk348b1.3 comes from this gene

cdna EST yk367e6.3 comes from this gene

```

CDNA EST yk236e4.5 comes from this gene
CDNA EST yk295d11.5 comes from this gene
CDNA EST yk348b1.5 comes from this gene
CDNA EST yk367e6.5 comes from this gene
CDNA EST yk313d8.3 comes from this gene
CDNA EST yk669d4.3 comes from this gene
CDNA EST yk671a8.3 comes from this gene
CDNA EST yk250f12.5 comes from this gene
CDNA EST yk513b8.5 comes from this gene
CDNA EST yk669p4.5 comes from this gene
CDNA EST yk671a8.5 comes from this gene
CDNA EST yk775g12.3 comes from this gene
CDNA EST yk775g12.5 comes from this gene
/codon_start=1
/product="Hypothetical protein R06c7.1"
/protein_id="CAA95839.1"
/db_xref="GI:3678834"
/db_xref="SPTREMBL:Q21770"
/translation="MSPHPPQHPMPMPVPTAPPGAMTPMPVPADAKLHQTGN
DACLRLQDLNVEDGAKMYMKPTPEKMKRPVDIOTNVGIEVTEKETTHERMNAKA
DLSTREKVFTEKKKEDFVODRDRDCCGIFLAEKNEPEFKMDQNOIVDGGSTL
YTTVNLFSFLDANGSKSKYQJNGADGDDDKLPCISLEIYARNDNTIISSENGS
KRTADONTEVNNREYTOLELALNQHCVNEIRFGCFERKGYFLNABEGDQRCV
DVGDGKOLYPLGLKTIQTEGPGYGRQNNPSLVIDGKMAFKEDQIVLOKLDITGD
PENGNNMTREKAAVVIKGLDQYSTYNNRRLRIEGLFHEBATIKRELPDQKCSI
AEVYADKYKISLOYPNANLVCKDRGNMYPAEIMYVSRNORVITPOOTGOSOKT
KECAVLPDVRQIITGNKAVNTLENNELVALGIKYVSEPLMVAORELDGKELYOR
SYMDSMGKWRAPGPFVPRVPTVDLMAAFAVNGPCRFISGVNOLYKAFIDSCKKG
MYIKPCEETGLSTERIKMTQLEKVAASKYVIMTDDIVLHKOYKALEORTMAYI
ODMKISKANAIVKDKRLTLENIINKTNVKGGLNTVSDARKSMTDEQLIGVQSA
PACRTKYMDNKNLPQIIGFASNAVNAHVEFGVFLAPSGDQTMASIEDVLONSID
LEKRNKALPKRIIIRYSGASESHASIIAYEIPLARAIHSEYSEIKLFIYVREH
SYRFRPDOLRSGKATENMIPPGIVLDNATVPACKOFLNHTTLOGTAKPIYVL
ADDKAPMDRELFTTLCCHHOIYVSIISIPPLVYANEAHKGRLDMLGELTTGPI
EAKESQGERLKEIKGKQYKODLNKRYNA"
gene
    complement(6739..11408)
    /gene="dhp-1"
    /complement(join(6739..6889,7899..7963,8094..8340,
8392..8457,8905..9182,9550..9761,9840..10017,10603..10783,
11317..11408))
    /gene="dhp-1"

Query Match      55.6%; Score 833; DB 3; Length 30911;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 448 AATTGCATTGATGCGCAAAATCGTGGCCAAACTTCGCTCTCCAAATTCGATGAGGG 507
    |||||||
Db 17800 AATTGCATTGATGCGCAAAATCGTGGCCAAACTTCGCTCTCCAAATTCGATGAGGG 17741

QY 508 AAGGCTCCTTAAGCAAGATCGTTCAAGATTGGCAACGCTTCAACTATTAAATTAN 567
    |||||||
Db 17740 AAGGCTCCTTAAGCAAGATCGTTCAAGATTGGCAACGCTTCAACTATTAAATTAN 17681

QY 568 TCCCAATTCTACTGAATACGCGTGAAGCGCAATTCAGAAATATGTGGAGCAGCAATGANT 627
    |||||||
Db 17680 TCCCAATTCTACTGAATACGCGTGAAGCGCAATTCAGAAATATGTGGAGCAGCAATGANT 17621

QY 628 GATCTATCTACAAAGAAAGACTTTCCCGAATGCTTCCAGATGACATGACGACAGACAA 687
    |||||||
Db 17620 GATCTATCTACAAAGAAAGACTTTCCCGAATGCTTCCAGATGACATGACGACAGACAA 17561

QY 688 GATCTTACGCTGATCTCAATATTTGATGAGAGCAGGAGGAGCTTCTCATATTCTCTGTT 747
    |||||||
Db 17560 GATCTTACGCTGATCTCAATATTTGATGAGAGCAGGAGGAGCTTCTCATATTCTCTGTT 17501

QY 748 GGATTTTGACAGCAGTCAATGATATCAACTAAATGCGAAAAAGCAATATATTGAGCAGACA 807
    |||||||
Db 17500 GGATTTTGACAGCAGTCAATGATATCAACTAAATGCGAAAAAGCAATATATTGAGCAGACA 17441

QY 808 AATTAATTTGCTCAAGCAATATAAATATGAGAAATATCCAAAGATATGATCTCAGACGAGTC 867
    |||||||
Db 17440 AATTAATTTGCTCAAGCAATATAAATATGAGAAATATCCAAAGATATGATCTCAGACGAGTC 17381
```

```

QY 868 ACATTGATCAATTATAGCAAAAGATCCCAATTGATCCCATGATTTGGAGAAAGTTAAGTT 927
    |||||||
Db 17380 ACATTGATCAATTATAGCAAAAGATCCCAATTGATCCCATGATTTGGAGAAAGTTAAGTT 17321

QY 928 GGACAAAAGTTTGAGCTCATGACGCCCTTGCTCAGCAATTCATTAACCTCCAGCTGCT 987
    |||||||
Db 17320 GGACAAAAGTTTGAGCTCATGACGCCCTTGCTCAGCAATTCATTAACCTCCAGCTGCT 17261

QY 988 TCGATTCTCAAAATTTTGGCAAAACTGAGAGATATCTTATGTGGGATGATGTCAGAT 1047
    |||||||
Db 17260 TCGATTCTCAAAATTTTGGCAAAACTGAGAGATATCTTATGTGGGATGATGTCAGAT 17201

QY 1048 GCACCTGAGACAGTTTCTCATTCATTCATTAATACATTTATGTTCCAGTGGTAT 1107
    |||||||
Db 17200 GCACCTGAGACAGTTTCTCATTCATTCATTAATACATTTATGTTCCAGTGGTAT 17141

QY 1108 GCGGAAAAGTATATTTGGAACCTGTCGCCGAGATGATCAAGAAACATTCAGATGG 1167
    |||||||
Db 17140 GCGGAAAAGTATATTTGGAACCTGTCGCCGAGATGATCAAGAAACATTCAGATGG 17081

QY 1168 GATGATCTCTGAGAAAGATCTGAGAAACCCCTGACCTGATCTGTTCAGACCAATG 1227
    |||||||
Db 17080 GATGATCTCTGAGAAAGATCTGAGAAACCCCTGACCTGATCTGTTCAGACCAATG 17021

QY 1228 CCTTCCCAAGAGATTTAGCAAAATTTAAGTATTTCTGATTTCCAAACGGGT 1280
    |||||||
Db 17020 CCTTCCCAAGAGATTTAGCAAAATTTAAGTATTTCTGATTTCCAAACGGGT 16968

RESULT 2
AC103229/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Rattus norvegicus
Rattus norvegicus (Norway rat)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 218764)
REFERENCE
AUTHORS
Muzny,D.,Marie., Metzker,M.,Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davilla,M.L., Davis,C., Davy-Carroll,L., De Arda,C., Deckerich,D.,
Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Dlyva,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Baves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebreygeorgis,E., Geer,K., Gilly,R., Grady,M., Guerra,M., Guevara,W.,
Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,J.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,L.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Huik,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolyet,A.,
Karpaty,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kover,C.,
Kowis,C., Kraft,C.L., Ledow,H., Levay,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Louisedge,H., Lozada,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindaratne,M., Mahmoud,M., Mallory,K., Mangum,A.,
Mangum,B., Mapa,P., Martin,K., McNeill,T.Z., Meenen,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Moore,S.,
Milosavljevic,A., Milner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munkidasa,M., Murphy,M., Nair,L.,
```

Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwokememh, O., Okwunu, G., Olarpunsaogon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C.,
 Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,
 Puzo, M., Quirroz, J., Rachlin, E., Reeves, K., Regier, M., Reigh, R.,
 Rellis, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.,
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Slason, I., Sletter, C., Sma, J.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
 Weinstein, G., and Gibbs, R. A.
 Direct Submission
 Title
 JOURNAL
 REFERENCE
 AUTHORS
 JOURNAL
 2 (bases 1 to 218764)
 Unpublished
 Direct Submission
 Title
 JOURNAL
 REFERENCE
 AUTHORS
 JOURNAL
 3 (bases 1 to 218764)
 Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On May 13, 2003 this sequence version replaced gi:22856257.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a contig-scaffold). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GUND
 Center clone name: CH230-217018
 ----- Summary Statistics
 Assembly program: Atlas 3.0:
 Consensus quality: 211841 bases at least Q40
 Consensus quality: 213096 bases at least Q30
 Consensus quality: 213870 bases at least Q20
 Estimated insert size: 218571; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

* 1 218764: contig of 218764 bp in length.
 Location/Qualifiers
 source
 1..218764
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-217018"
 1..1339
 /note="wgs-contig"
 BASE COUNT 63585 a 42997 c 43704 g 64011 t 4467 others
 ORIGIN
 Query Match 1.7%: Score 25; DB 2; Length 218764;
 Best Local Similarity 100.0%; Pred. No. 0.068;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 5 CTGAATTTCGAAATTCACAGC 29
 |||||||||||||||||||||
 Db 74190 CTGAATTTCGAAATTCACAGC 74166
 RESULT 3
 AP005398/C 133343 bp DNA linear HTG 07-JUN-2002
 LOCUS
 DEFINITION Oryza sativa (japonica cultivar-group) chromosome 2 clone p0669G10,
 *** SEQUENCING IN PROGRESS ***.
 ACCESSION
 AP005398
 VERSION
 AP005398.1 GI:21328228
 KEYWORDS
 HTG; HTGS; PHASE2.
 SOURCE
 Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Euphorbiaceae; Oryzae; Oryza.
 REFERENCE
 1
 Sasaki, T., Matsumoto, T. and Katayose, Y.
 Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, PAC
 clone:p0669G10
 Published Only in Database (2002)
 2 (bases 1 to 133343)
 Sasaki, T., Matsumoto, T. and Katayose, Y.
 Direct Submission
 Submitted (06-JUN-2002) Takuji Sasaki, National Institute of
 Agrobiological Sciences, Rice Genome Research Program, Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail: tsasaki@nias.affrc.go.jp, UR: <http://rgp.dna.affrc.go.jp/>,
 Tel: 81-298-38-7441, Fax: 81-298-38-7468)
 NOTE: It currently consists of 1 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces is believed
 to be correct as given, however the sizes of the gaps between them
 are based on estimates that have provided by the submitter. This
 sequence will be replaced by the finished sequence as soon as it is
 available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 Location/Qualifiers
 1..133343
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="2"
 /clone="p0669G10"
 BASE COUNT 38280 a 26701 c 27851 g 40458 t 53 others
 ORIGIN
 Query Match 1.5%; Score 23; DB 2; Length 133343;
 Best Local Similarity 100.0%; Pred. No. 0.92;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 824 CATTAAATAATGAGAAATTCGA 846

DB 4277 CAATAAAATGAGAAATCCA 4255

RESULT 4 AC091791 LOCUS

AC091791 157654 bp DNA linear HTG 19-APR-2002
Sus scrofa clone RP44-497D19, WORKING DRAFT SEQUENCE, 3 ordered
pieces.

AC091791
AC091791.2 GI:20198527
HTG: HTGS_PHASE2; HTGS_DRAFT.
Sus scrofa (pig)
Sus scrofa

REFERENCE AUTHORS

1 (bases 1 to 157654)
Akhter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C.,
Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J.,
Haghighi, P., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Laric, P.,
Lee-Lin, S.-Q., Legaspi, R., Maduro, Q.L., Maduro, V.B.,
Marquies, E.H., Masello, C., Maskeri, B., Mestrilan, S.D.,
McCloskey, J.C., McDowell, J., Paquirigan, C., Pearson, R.,
Portnoy, M.E., Prasad, A., Schueler, M.G., Stanciripop, S., Thomas, J.W.,
Thomas, P.J., Touchman, J.W., Tsurgou, C., Vogt, J.L., Walker, M.A.,
Welter, P.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
NISC Comparative Sequencing Initiative

Unpublished
2 (bases 1 to 157654)

REFERENCE AUTHORS

Green, E.D.
Direct Submission
Submitted (07-JUN-2001) NIH Intramural Sequencing Center, 8717
Groveport Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 157654)

REFERENCE AUTHORS

Green, E.D.
Direct Submission
Submitted (19-APR-2002) NIH Intramural Sequencing Center, 8717
Groveport Circle, Gaithersburg, MD 20877, USA
On Apr 19, 2002 this sequence version replaced gi:14327784.

COMMENT

Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc.zoo@nhgri.nih.gov
----- Project Information
Center project name: caf
Center clone name: 497D19

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones; overlaps with neighboring clones; alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 157132 bases at least Q40
Consensus quality: 157327 bases at least Q40
Consensus quality: 157401 bases at least Q20
Insert size: 154000; agarose-fp
Insert size: 157454; sum-of-contigs
Quality coverage: 12.93x in Q20 bases; agarose-fp
Quality coverage: 12.64x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.
* 1 128692: contig of 128692 bp in length
* 128693 128792: gap of unknown length
* 128793 144829: contig of 16037 bp in length
* 144830 144929: gap of unknown length
* 144930 157654: contig of 12725 bp in length.
Location/Qualifiers
1.157654
/organism="Sus scrofa"
/mol_type="genomic DNA"
/db_xref="taxon:9823"
/clone="RP44-497D19"
/clone_lib="RP44"
1.128692
/note="assembly_fragment"
/note="assembly_fragment"
vector_side:left"
1.92843
/note="clone overlaps with Genbank Accession Number AC091790 clone RP44-278652 (center project name cae)"

misc_feature
1.128692
/note="assembly_fragment"
vector_side:left"
1.92843
/note="clone overlaps with Genbank Accession Number AC091790 clone RP44-278652 (center project name cae)"
misc_feature
128793..144829
/note="assembly_fragment"
129387..157654
/note="clone overlaps with Genbank Accession Number AC091722 clone RP44-304F21 (center project name akv)"
144930..157654
/note="assembly_fragment"
clone_end:SP6
vector_side:right"

BASE COUNT 41793 a 33973 c 35855 g 45833 t 200 others
ORIGIN

Query Match 1.5%; Score 22; DB 2; Length 157654;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 660 GCTTCAGATGCGATGACGAC 681
DB 7180 GCTTCAGATGCGATGACGAC 7201

RESULT 5 AC128697/c

LOCUS AC128697 183239 bp DNA linear HTG 25-AUG-2002
DEFINITION Mus musculus chromosome UNK clone RP24-388P9, WORKING DRAFT
SEQUENCE, 9 unordered pieces.

AC128697
AC128697.2 GI:22476247
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Mus musculus (house mouse)
Mus musculus

REFERENCE AUTHORS

McPherson, J.D. and Waterston, R.H.
The sequence of Mus musculus clone
Unpublished
2 (bases 1 to 183239)
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (22-JUN-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 183239)
McPherson, J.D. and Waterston, R.H.
Direct Submission

JOURNAL Submitted (25-AUG-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
COMMENT On Aug 25, 2002 this sequence version replaced g1:21914538.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
Contact: submissions@wustl.wustl.edu
Project Information
Center project name: M.B80388P09

----- Summary Statistics -----

Sequencing vector: M13; 08
Sequencing vector: plasmid: 1008
Chemistry: Dye-terminator ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 177227 bases at least Q40
Consensus quality: 178348 bases at least Q30
Consensus quality: 178997 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2586: contig of 2586 bp in length
* 2687 2686: gap of unknown length
* 5770 5769: gap of 3083 bp in length
* 5870 5869: gap of unknown length
* 12438 12437: contig of 6568 bp in length
* 12538 12537: gap of unknown length
* 34519 34518: contig of 21981 bp in length
* 34619 34618: gap of unknown length
* 53627 53626: contig of 19008 bp in length
* 53727 53726: gap of unknown length
* 75557 75556: contig of 21830 bp in length
* 75657 75656: gap of unknown length
* 100968 100967: contig of 25311 bp in length
* 101068 101067: gap of unknown length
* 131088 131087: contig of 30021 bp in length
* 131189 131188: gap of unknown length
* 131189 183239: contig of 52051 bp in length.

FEATURES

source

1. 183239
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="X" "X"
/clone="RP24-388P9"
1. 2586
/note="assembly_name:Contig9"
misc_feature 2687..5769
/note="assembly_name:Contig10"
misc_feature 5870..12437
/note="assembly_name:Contig11"
misc_feature 12538..34518
/note="assembly_name:Contig12"
misc_feature 34619..53626
/note="assembly_name:Contig13"
misc_feature 53727..75556
/note="assembly_name:Contig14"
misc_feature 75657..100967
/note="assembly_name:Contig15"
misc_feature 101068..131088
/note="assembly_name:Contig16"
misc_feature 131189..183239
/note="assembly_name:Contig17"
BASE COUNT 51407 a 40699 c 39395 g 50857 t 881 others

ORIGIN

Query Match 1.5%; Score 22; DB 2; Length 183239;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 214 AACTGCATTCGATTCAGTCA 235
|||||
Db 153697 AACTGCATTCGATTCAGTCA 153676

RESULT 6

AC118694 193311 bp DNA linear HTG 23-APR-2003
LOCUS Mus musculus clone RP24-406H23, WORKING DRAFT SEQUENCE, 6 unordered
DEFINITION pieces.

AC118694
AC118694.6 GI:30018084
VERSION HTG; HTGS; PHASE1; HTGS_DRAFT.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus

REFERENCE

AUTHORS

1 (bases 1 to 193311)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, A., Allen, N.,
Anderson, S., Brown, A., Cantarel, D., Chang, J.,
Boukhalter, B., Brown, A., Cantarel, D., Chang, J.,
Chazaro, B., Choquet, Y., Colangelo, M., Collins, S., Collins, S.,
Cook, A., Cooke, P., DeArrelano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Larocque, R., Lamazares, R.,
Lander, T., Lebeck, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menus, L.,
Mihova, T., Mleong, V., Murphy, T., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

REFERENCE

AUTHORS

2 (bases 1 to 193311)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, A., Allen, N.,
Anderson, S., Brown, A., Cantarel, D., Chang, J.,
Boukhalter, B., Brown, A., Cantarel, D., Chang, J.,
Chazaro, B., Choquet, Y., Colangelo, M., Collins, S., Collins, S.,
Cook, A., Cooke, P., DeArrelano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Larocque, R., Lamazares, R.,
Lander, T., Lebeck, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menus, L.,
Mihova, T., Mleong, V., Murphy, T., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

JOURNAL

Submitted (20-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 193311)

REFERENCE

AUTHORS

Birren, B., Nusbaum, C., Lander, E., Abouelell, A., Allen, N.,
Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choquet, Y.,
Collins, S., Cooke, P., Cooke, P., Corum, B., DeArrelano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Lander, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Liu, G., Mabbitt, R., Maclean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrum, J., Menus, L., Mihova, T., Mleong, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,

TITLE
JOURNAL
COMMENT

Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-APR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 17, 2003 this sequence version replaced gi:28467223.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L21128
Center clone name: 406_H23

* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 4849: contig of 4849 bp in length
* 4850 4949: gap of 100 bp
* 4950 99766: contig of 94817 bp in length
* 99767 99866: gap of 100 bp
* 99867 135133: contig of 35267 bp in length
* 135134 135233: gap of 100 bp
* 135234 156045: contig of 20812 bp in length
* 156046 156145: gap of 100 bp
* 156146 184734: contig of 28589 bp in length
* 184735 184834: gap of 100 bp
* 184835 193311: contig of 8477 bp in length.

FEATURES
source
1.193311
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP24-406H23"
/clone_1lb="RPC1-24 Male Mouse BAC"
1.4849
/note="assembly-fragment
clone_end:SP6
vector_side:left"
misc_feature
4950..99766
/note="assembly-fragment"
99867..135133
/note="assembly-fragment"
135234..156045
/note="assembly-fragment"
156146..184734
/note="assembly-fragment"
184835..193311
/note="assembly-fragment
clone_end:T7
vector_side:right"
BASE COUNT 54331 a 40998 c 43020 g 54457 t 505 others
ORIGIN

Query Match 1.5%; Score 22; DB 2; Length 193311;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 AACTGGATTCGATCAAGTCA 235
|||||
Db 41350 AACTGGATTCGATCAAGTCA 41371

RESULT 7
AC091790

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AC091790 195037 bp DNA linear HTG 12-APR-2002
Sus scrofa clone RP44-278G22, WORKING DRAFT SEQUENCE, 4 ordered
pieces.
AC091790
AC091790.2 GI:20143538
HTG: HTGS_PHASE2: HTGS_DRAFT.
Sus scrofa (pig)
Sus scrofa

REFERENCE
AUTHORS

1 (bases 1 to 195037)
Akheri,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,
Brooks,S., Dietrich,N.L., Grantle,S., Guan,X., Gupta,J.,
Haghighi,P., Hansen,N., Ho,S.-L., Ido,J.R., Karlins,E., Laric,P.,
Lee-Uin,S.-Q., Legaspi,R., Maduro,O.L., Maduro,V.B.,
Margulies,E.H., Mastello,C., Maskeri,B., Mastrian,S.D.,
McCloskey,J.C., McDowell,J., Pagurigan,C., Pearson,R.,
Portnoy,M.E., Prasad,A., Schueler,M.G., Stantirlop,S., Thomas,J.W.,
Thomas,P.J., Touchman,J.W., Tsurgoun,C., Vogt,J.L., Walker,M.A.,
Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 195037)
Green,E.D.
Direct Submission
Submitted (07-JUN-2001) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 195037)
Green,E.D.
Direct Submission
Submitted (12-APR-2002) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA
On Apr 12, 2002 this sequence version replaced gi:14327783.

COMMENT

----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nhgri.nih.gov
----- Project Information
Center project name: cae
Center clone name: 278G22

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8X average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 193528 bases at least Q40
Consensus quality: 194261 bases at least Q30
Consensus quality: 194570 bases at least Q20
Insert size: 144000; agarose-fp
Insert size: 194737; sum-of-contigs
Quality coverage: 11.89x in Q20 bases; sum-of-contigs
Quality coverage: 8.79x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.

```
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 14284: contig of 14284 bp in length
* 14285 26042: contig of 11658 bp in length
* 14385 26042: gap of unknown length
* 26043 26142: gap of unknown length
* 26143 64611: contig of 38469 bp in length
* 64612 64711: gap of unknown length
* 64712 195037: contig of 130326 bp in length.
Location/Qualifiers
1. 195037
/organism="Sus scrofa"
/mol_type="genomic DNA"
/db_xref="taxon:9823"
/clone="RP44-278G22"
/clone_11b="RP44"
misc-feature
1. 45776
/note="clone overlaps with GenBank Accession Number
AC092194 clone RP44-309D18 (center project name cad)"
misc-feature
1. 14284
/note="assembly_fragment
clone_end:17
vector_side:left"
misc-feature
14385..26042
/note="assembly_fragment"
26143..64611
/note="assembly_fragment"
64712..195037
/note="assembly_fragment
clone_end:SP6
vector_side:right"
102196..195037
/note="clone overlaps with GenBank Accession Number
AC091791 clone RP44-497D19 (center project name caf)"
BASE COUNT 51409 a 42593 c 44365 g 56370 t 300 others
ORIGIN
Query Match 1.5%; Score 22; DB 2; Length 195037;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 660 GCTTCCAGATGACATGACGAC 681
|||||
Db 109375 GCTTCCAGATGACATGACGAC 109396

RESULT 8
AC140183 205682 bp DNA linear HTG 23-FEB-2003
LOCUS Mus musculus chromosome UNK clone RP24-350N4, WORKING DRAFT
DEFINITION
SEQUENCE, 3 unordered pieces.
AC140183
AC140183.1 GI:28475416
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 205682)
McPherson,J.D. and Waterston,R.H.
The sequence of Mus musculus clone
Unpublished
2 (bases 1 to 205682)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (23-FEB-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
```

```
Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: M_B80350N04
----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 204725 bases at least Q40
Consensus quality: 204994 bases at least Q30
Consensus quality: 205269 bases at least Q20
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 38205: contig of 38205 bp in length
* 38206 38305: gap of unknown length
* 38306 99264: contig of 60959 bp in length
* 99265 99365: gap of unknown length
* 99365 205682: contig of 106318 bp in length.
Location/Qualifiers
1. 205682
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="XNK"
/clone="RP24-350N4"
misc-feature
1. 38205
/note="assembly_name:Contig15"
38306..99264
/note="assembly_name:Contig16"
99365..205682
/note="assembly_name:Contig17"
BASE COUNT 64084 a 45447 c 42625 g 53326 t 200 others
ORIGIN
Query Match 1.5%; Score 22; DB 2; Length 205682;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1248 CAAATTTAAGGTAAATTCGATT 1269
|||||
Db 67895 CAAATTTAAGGTAAATTCGATT 67916

RESULT 9
AC118916 229331 bp DNA linear HTG 19-NOV-2002
LOCUS Rattus norvegicus clone CH230-177J9, WORKING DRAFT SEQUENCE.
DEFINITION
AC118916
AC118916.4 GI:25090461
HTG: HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 229331)
Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blais,J., Blankenburg,K., Blyth,P., Brown,M.,
Byrant,N., Buhay,C., Burch,P., Butrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
```

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunatille, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Haylak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hoggins, J., Hollins, B., Howells, S., Huik, S., Hume, J., Idlebird, D., Jackson, A., Jackson, B., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensunewa, L., Louisedge, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Mallory, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Millaavljovic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munitasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoelameh, O., Okwuonu, G., Olarinmusa, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reih, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, R., Ruiz, S.J., Sanders, M., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlaczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, J., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 229331)
Worley, K.C.

Direct Submission
Submitted (21-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 229331)

Rat Genome Sequencing Consortium.
Direct Submission
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 19, 2002 this sequence version replaced gi:23265446.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center Project name: GWXU
Center Clone name: CH230-17739
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 208116 bases at least Q40
Consensus quality: 209830 bases at least Q30
Consensus quality: 211071 bases at least Q20
Estimated insert size: 210933; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 229331: contig of 229331 bp in length.
Location/Qualifiers
1. 229331
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-17739"
1. 1213
/note="wgs_end_extension
/clone_end:Sp6
7353. 8196
/note="clone_boundary
clone_end:Sp6
site:
end_sequence: BH274141"
complement(223960..224162)
/note="clone_boundary
clone_end:T7
site:
end_sequence: BH274138"
226389..227418
/note="wgs_end_extension
clone_end:T7"
227469..229331
/note="wgs_end_extension
clone_end:T7"
BASE COUNT 58707 a 45353 c 47288 g 60852 t 17151 others
ORIGIN
Query Match 1.5%; Score 22; DB 2; Length 229331;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1387 GATGAGAAATTGATGACTGT 1408
Db 134531 GATGAGAAATTGATGACTGT 134552

RESULT 10
AC094981/C 231442 bp DNA 1linear HTG 09-MAY-2003
LOCUS AC094981
DEFINITION Rattus norvegicus clone CH230-6121, WORKING DRAFT SEQUENCE.
ACCESSION AC094981
VERSION AC094981.6 GI:30467284
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 231442)

AUTHORS
 Muzny, D., Marle, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angiano, D., Anyalebechi, V., Ayogbi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, R., Blair, D., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, E., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Friser, C., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gargregegis, E., Geer, K., Gill, R., Grady, M., Guerra, N., Guerrero, W., Gunaratne, P., Healand, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlik, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S., L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kwis, C., Kraft, C., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lotensuhewa, L., Louisedge, H., Lozano, R., Lu, X., Ma, J., Maheshwari, M., Mahindartine, M., Mahmoud, M., Malloy, K., Mangun, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M., McNeill, T., Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Morgan, M., Morris, K., Morris, S., Munnidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokkelumeh, O., Okwunonu, G., Olariunpagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C., Plummer, F., Polidexter, A., Popovic, D., Primus, E., Pu, L., L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M., A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Slisson, I., Sitter, C., D., Smas, J., Sneed, A., Sodergren, E., Song, X., Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Vallas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, S., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D., R., Holt, R., A., Smith, H., O., Weinstein, G., and Gibbs, R. A.

TITLE
 Direct Submission
 Unpublished
 2 (bases 1 to 231442)

JOURNAL
 Direct Submission
 Worley, K. C.

REFERENCE
 Submitted (15-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 231442)
 Rat Genome Sequencing Consortium.

AUTHORS
 Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT
 On May 9, 2003 this sequence version replaced g1:22771413. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole

genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine

Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information

Center project name: GBMC
 Center clone name: CH230-6121

----- Summary Statistics

Assembly program: Atlas;

Consensus quality: 207859 bases at least Q40
 Consensus quality: 211922 bases at least Q30

Consensus quality: 214681 bases at least Q20
 Estimated insert size: 220165; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have
 * provided by the submitter.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and

* the accession number will be preserved.

1 231442: contig of 231442 bp in length.

FEATURES

source

1. 231442

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-6121"

1. 1860

/note="wgs_end_extension"

clone_end:77

5888..6759

/note="clone_boundary"

clone_end:77

site:ECORI

end_sequence:BH360141"

229224..229921

/note="clone_boundary"

clone_end:Sp6

site:ECORI

end_sequence:BH360143"

BASE COUNT 63051 a 48424 c 46447 g 58733 t 14787 others

ORIGIN

Query Match 1.5%; Score 22; DB 2; Length 231442;

Best Local Similarity 100.0%; Pred. No. 3.6;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

misc-feature

Db 72367 AACTCGATTGCATTCAAGTCA 72346

RESULT 11
 AC128235 253961 bp DNA linear HTG 21-SEP-2002
 AC128235/c
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-38G15, *** SEQUENCING IN PROGRESS
 *** 4 unordered pieces.
 AC128235
 AC128235.2 GI:23265191
 VERSION
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae:
Rattus.
1 (bases 1 to 253961)

REFERENCE
AUTHORS
Muzny,D,Marie, Metzker,M,lee, Abramson,S, Adams,C, Alder,J, Allen,C, Allen,H, Alsprouks,S, Amin,A, Anguiano,D, Anyalebech,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H, Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F, Bisvalo,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M, Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E, Cardenas,V, Carter,K, Cavazos,I, Caesar,H, Center,A, Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J, Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,D, Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D, Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Divya,K, Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K, Egan,A, Escotto,M, Eugene,C, Evans,C, Falls,T, Fan,G, Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M, Gebregeorgis,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,W, Gunaratne,P, Haaland,W, Hamill,C, Hamilton,C, Hamilton,K, Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J, Hernandez,R, Hines,S, Hulyk,S, Hume,J, Idlebird,D, Jackson,A, Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A, Karpathy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C, Kowis,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,D, Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J, Lorensbawa,L, Loulseged,H, Lozado,R,J, Lu,X, Ma,J, Maheshwari,M, Mahindaratne,M, Mahmoud,M, Malloy,K, Mangum,A, Mangum,B, Mapua,P, Martin,K, Martin,R, Martinez,E, Mawhney,S, McLeod,M,P, McNeill,T,Z, Meenen,E, Milsavajevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S, Morris,K, Morris,S, Munitasa,M, Murphy,M, Natl,L, Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S, Nwaekelemel,O, Okwunu,G, Olarunpusogun,A, Pal,S, Parks,K, Pasternak,S, Paul,H, Perez,A, Perez,L, Pfankoch,C, Plopper,F, Poindexter,A, Popovic,D, Primus,E, Pu,L,L, Puazo,M, Quiroz,J, Rachlin,E, Reeves,K, Regier,M,A, Relgh,R, Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F, Rives,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,J, Sanders,W, Savery,G, Scherer,S, Scott,G, Shatsman,S, Shen,H, Shetty,J, Shvartsbeyn,A, Slisson,I, Sitter,C,D, Smajs,D, Sneed,A, Sodergren,E, Song,X,Z, Sorelle,R, Sosa,J, Steimle,M, Strong,R, Sutton,A, Svatek,A, Taber,P, Taylor,C, Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Usmani,K, Valas,R, Vera,V, Villasana,D, Waldron,L, Walker,B, Wang,J, Wang,Q, Wang,S, Warren,J, Warren,R, Wei,X, White,F, Williams,G, Willson,R, Wleczyk,R, Wooden,H, Worley,K, Wright,D, Wright,R, Wu,J, Yakub,S, Yen,J, Yoon,L, Yoon,V, Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Dunn,D, Von Niederhausern,A, Weiss,R, Smith,D,R, Holt,R,A, Smith,H,O, Weinstock,G, and Gibbs,R.A.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
2 (bases 1 to 253961)
Worley,K.C.
Direct Submission
Submitted (19-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 253961)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (21-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Sep 21, 2002 this sequence version replaced gi:21908845.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only

contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GYBV

Center clone name: CH230-38G15

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 214169 bases at least Q40

Consensus quality: 218921 bases at least Q30

Consensus quality: 222222 bases at least Q20

Estimated insert size: 237536; sum-of-contigs estimation

Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 4 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 249583: contig of 249583 bp in length

* 249584 249683: gap of unknown length

* 249684 250819: contig of 1136 bp in length

* 250820 250919: gap of unknown length

* 250920 252666: contig of 1747 bp in length

* 252667 252766: gap of unknown length

* 252767 253961: contig of 1195 bp in length.

Location/Qualifiers

1. 253961

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-38G15"

1. 1470

/note="wgs_end-extension

clone_end:77"

4785. 5385

/note="clone boundary

clone_end:77

site:ECORI

end.sequence:BH277815"

174876. 175881

/note="wgs_contig"

177111. 178416

/note="wgs_contig"

BASE COUNT 63786 a 51180 c 48941 g 61233 t 28821 others

ORIGIN

Query Match

Best local similarity 100.0%; Pred. No. 3.7;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 AACTGGATTCGATTCACGCTCA 235

|||||

DB 217452 AACTGGATTCGATTCACGCTCA 217431

RESULT 12

BD146582

LOCUS

BD146582 872 bp DNA linear PAT 17-JAN-2003

DEFINITION Primer for synthesizing full-length cDNA and use thereof.

ACCESSION BD146582

BD146582.1 GI:27852340

KEYWORDS UP 2002191363-A/1425.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

```

REFERENCE
AUTHORS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE        1 (bases 1 to 872)
JOURNAL      Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
              Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
              Primer for synthesizing full-length cDNA and use thereof
              Patent: JP 2002191363-A 1425 09-JUL-2002;
              HELIX RESEARCH INSTITUTE
COMMENT      OS Homo sapiens (human)
              PN JP 2002191363-A/1425
              PD 09-JUL-2002
              PF 28-JUL-2000 JP 2000280990
              PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
              PI SAITO,
              PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
              PI KEIICHI NAGAI,TETSUJI OTSUKI
              PC
C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/
10,
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH key
Location/Qualifiers
FT source 1..872
          /organism="Homo sapiens (human)".
FEATURES
Source    1..872
          Location/Qualifiers
          1..872
          /organism="Homo sapiens"
          /mol_type="genomic DNA"
          /db_xref="taxon:9606"
BASE COUNT      285 a 138 c 177 g 263 t 9 others
ORIGIN
Query Match      1.4%; Score 21; DB 6; Length 872;
Best Local Similarity 100.0%; Pred. No. 8.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      873 TGATCAATTAGCAAAAGATCC 893
DB      747 TGATCAATTAGCAAAAGATCC 767

RESULT 13
BD159393
LOCUS      BD159393      2202 bp      DNA      linear      PAT 17-JAN-2003
DEFINITION      Primer for synthesizing full-length cDNA and use thereof.
ACCESSION      BD159393.1 GI:27865151
VERSION      BD159393.1
KEYWORDS      JP 2002191363-A/14236.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2202)
Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 14236 09-JUL-2002;
HELIX RESEARCH INSTITUTE
COMMENT      OS Homo sapiens (human)
              PN JP 2002191363-A/14236
              PD 09-JUL-2002
              PF 28-JUL-2000 JP 2000280990
              PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
              PI SAITO,
              PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
              PI KEIICHI NAGAI,TETSUJI OTSUKI
              PC
C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/
10,
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH key
Location/Qualifiers
FT source 1..2202

```

```

FEATURES
Source    Location/Qualifiers
          1..2202
          /organism="Homo sapiens (human)".
          /mol_type="genomic DNA"
          /db_xref="taxon:9606"
BASE COUNT      703 a 416 c 488 g 595 t
ORIGIN
Query Match      1.4%; Score 21; DB 6; Length 2202;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      873 TGATCAATTAGCAAAAGATCC 893
DB      746 TGATCAATTAGCAAAAGATCC 766

RESULT 14
AK022163
LOCUS      AK022163      2202 bp      mRNA      linear      PRI 01-AUG-2002
DEFINITION      Homo sapiens cDNA FLJ12101 fis, clone HEMBB1002683.
ACCESSION      AK022163
VERSION      AK022163.1 GI:10433497
KEYWORDS      cDNA capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2202)
Ota,T., Isogai,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Takahashi-Fujii,A., Hara,H.,
Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K.,
Arita,M., Nebekura,T., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J.,
Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuno,Y. and Oshima,A.
NEDO human cDNA sequencing project
Unpublished
JOURNAL      2 (bases 1 to 2202)
AUTHORS      Isogai,T. and Otsuki,T.
TITLE        Direct Submission
JOURNAL      Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,
              Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
              (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
              NEDO human cDNA sequencing project supported by Ministry of
              International Trade and Industry of Japan; cDNA full insert
              sequencing: Research Association for Biotechnology; cDNA library
              construction; 5'- & 3'-end one pass sequencing and clone selection:
              Helix Research Institute (supported by Japan Key Technology Center
              etc.) and Department of Virology, Institute of Medical Science,
              University of Tokyo.
              Location/Qualifiers
              1..2202
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="HEMBB1002683"
              /issue_type="whole embryo, mainly body"
              /clone_lib="HEMBB1"
              /dev_stage="embryo, 10 weeks"
              /note="cloning vector: pME18SFL3"
BASE COUNT      703 a 416 c 488 g 595 t
ORIGIN
Query Match      1.4%; Score 21; DB 9; Length 2202;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      873 TGATCAATTAGCAAAAGATCC 893
DB      746 TGATCAATTAGCAAAAGATCC 766

RESULT 15

```


DMC137E7/c 11922 bp DNA linear INV 27-APR-1999
 LOCUS DMC137E7
 DEFINITION Drosophila melanogaster cosmid clone 137E7.
 ACCESSION AL021108
 VERSION AL021108.1 GI:4164288
 KEYWORDS
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 11922)
 Ferraz,C., Vidal,S., Brun,C., Bucheton,A. and Demaille,J.G.
 Sequencing the distal X chromosome of Drosophila melanogaster
 Unpublished
 Institut de Genetique Humaine -UPR 1142- CNRS, 141,rue de la
 Cardonille, 34396 Montpellier Cedex 5, France
 2 (bases 1 to 11922)
 Benos,P.
 Direct Submission
 Submitted (26-APR-1999) European Drosophila Genome Sequencing
 Consortium
 On Jan 19, 1999 this sequence version replaced gi:3947671.
 Sequence submitted by Takis Benos, EMBL Outstation - The EBI,
 Hinxton, Cambridge, CB10 1SD, U.K.
 E-mail: benoseebi.ac.uk on behalf of the European Drosophila Genome
 Sequencing Consortium. For further information see the European
 Drosophila Genome Sequencing Consortium's web site:
 http://edgp.ebi.ac.uk/.
 Coding sequences are predicted from computer analysis, using both
 gene and CDS prediction programs and matches to other sequences.
 These predictions and matches have been evaluated by the annotators
 and may have been refined by hand (in which case a GeneFinder
 prediction will have no score. The annotators have also used their
 judgement on what matches to represent in this record. A far more
 complete annotation record is available from FlyBase
 (http://flybase.bio.indiana.edu/) through the FlyBase Annotation
 Object linked by the db_xref qualifier in the feature table.
 IMPORTANT: This sequence is NOT necessarily the entire insert of
 clone 137E7. It may be shorter, since we are minimising the
 overlap between clones to 100 bases, by trimming them. Clone 137E7
 overlaps to the left with clone 30B7
 The true left end of clone 137E7 is at position 4266 of clone 30B7
 Clone 137E7 overlaps to the right with clone 131F2 Sequence in
 absolute orientation with respect to chromosome The syntax for the
 representation of annotation used in this record is documented at:
 ftp://ftp.ebi.ac.uk/pub/databases/edgp/sequence_annotation.README
 Clone-137E7; Contig ID=1; Length=11922; Status=Finished.
 Location/Qualifiers
 1..11922
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="cosmid 137E7"
 4333..4524
 /gene="EG:137E7.1"
 4333..4524
 /note="EG:137E7.1"
 /note="prediction(method:'genefinder',
 version:'084', score:'12.36')"
 /prediction=(method:'genscan', version:'1.0')
 /match(desc:'LD19625.5prime LD Drosophila melanogaster
 embryo Bluescript Drosophila melanogaster cDNA clone
 LD19625 5prime, mRNA sequence', species:'Drosophila
 melanogaster (fruit fly)', ranges:(query:3905..4079,
 target:EMBL:AA540213:1..175, score:'848.00'))
 method:'blastn', version:'1.4.9')
 the EST EMBL:AA540213 comes from the 5' UTR"
 /codon_start=1
 /protein_id="CAA15947.1"
 /db_xref="GI:2760070"
 /db_xref="FLYBASE:FBgn0023543"
 /db_xref="SPTRMBL:O46049"

/translation="MEQITNYFEHNDYVSFSVALPCVVIANYGFLQVASGLACKSK
 SRMPGVCFHLTNQRIAD"
 BASE COUNT 3384 a 2429 c 2431 g 3678 t
 ORIGIN
 Query Match 1.4%; Score 21; DB 3; Length 11922;
 Best Local Similarity 100.0%; Pred.No. 11;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 409 AACGAAGATACCAAAATGAT 429
 ||||||||||||||||||||
 DB 9602 AACGAAGATACCAAAATGAT 9582
 Search completed: September 2, 2003, 19:33:31
 Job time : 5691 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 2, 2003, 20:31:10 : Search time 3066 seconds

(without alignments)
3947.694 Million cell updates/sec

Title: US-09-872-523-5
Perfect score: 498
Sequence: 1 MSEFLKIVRANKSDRKLDK.....PIGWCEANSYLVQPPKKNY 498

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Word size: 1

Total number of hits satisfying chosen parameters: 45562604

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODER=frame_plus_p2n.model -DEV=x1h
-Q/cgn2.1/USPTO.spool/US09872523/runat.02092003.073002.27259/app.query.fasta.1.647
-DB-EST -QFMT=fastap -SUFFIX=ol1.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM-ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09872523 @CGN1.1.2810 @runat.02092003.073002.27259 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEDUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC
-DEV_TIMECUT=120 -WARN_TIMECUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estinu:*
5: em_estlov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inu:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	210	42.2	653	9	AU204939
2	186	37.3	581	12	Bj101701
3	164	32.9	557	12	Bj124772
4	148	29.7	519	12	Bj115564
5	138	27.7	768	12	Bj143186
6	135	27.1	761	12	Bj129361
7	128	25.7	607	14	CB404801
8	119	23.9	360	14	C42305
9	112	22.5	380	9	AV197399
10	111	22.3	677	9	AU216908
11	108	21.7	686	12	Bj153087
12	102	20.5	340	9	AU111588
13	72	14.5	300	9	AU116296
14	70	14.1	300	9	AU115428
15	62	12.4	337	9	AU112450
16	22	4.4	505	14	CB385690
17	10	2.0	504	28	AO605493
18	9	1.8	196	29	AG241054
19	9	1.8	391	29	AG229361
20	9	1.8	456	12	Bj097227
21	9	1.8	494	12	Bj097088
22	9	1.8	502	10	BG668175
23	9	1.8	534	28	AO415204
24	9	1.8	608	12	Bj097296
25	9	1.8	610	12	Bj097211
26	9	1.8	626	12	Bj097047
27	9	1.8	637	12	Bj097165
28	9	1.8	644	12	Bj097081
29	9	1.8	655	28	AO389994
30	9	1.8	663	12	AO389984
31	9	1.8	693	12	Bj096984
32	9	1.8	701	28	Bj1290
33	9	1.8	704	12	Bj096997
34	9	1.8	715	29	AG071711
35	9	1.8	722	14	CB445305
36	9	1.8	738	14	CB958090
37	9	1.8	739	14	CB445578
38	9	1.8	793	12	Bj096871
39	9	1.8	877	14	CD303479
40	9	1.8	917	14	CD359252
41	8	1.6	93	9	Aj1206304
42	8	1.6	184	28	B64089
43	8	1.6	188	9	Aj155920
44	8	1.6	190	9	AA593206
45	8	1.6	190	28	AO081464

ALIGNMENTS

RESULT 1
LOCUS AU204939 653 bp mRNA linear EST 17-JUL-2001
DEFINITION AU204939 unpublished oligo-capped cDNA library, stage 1A
ACCESSION AU204939
VERSION AU204939.1 GI:14836792
KEYWORDS
SOURCE
ORGANISM
EST.
Caenorhabditis elegans
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea
; Rhabditidae; Peioderinae; Caenorhabditis.
REFERENCE
1 (bases 1 to 653)

AUTHORS Kohara, Y., Shin-I, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y. and Sugano, S.
 TITLE A complementary view of the C.elegans genome
 JOURNAL Unpublished
 COMMENT Contact: Yui Kohara
 Genome Biology Lab.
 National Institute of Genetics
 Yata 1111, Mishima, Shizuoka 411, Japan
 Tel: 81-559-81-6854
 Fax: 81-559-81-6855
 Email: ykohara@nig.ac.jp.

FEATURES
 source Location/Qualifiers

1..653
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /strain="N2"
 /db_xref="taxon:6239"
 /clone="YK843e05"
 /sex="Hermaphrodite"
 /tissue_type="whole animal"
 /dev_stage="L4"
 /clone_lib="unpublished oligo-capped cDNA library, stage L4"

BASE COUNT 216 a 116 c 136 g 185 t
 ORIGIN

Alignment Scores:

Score: 4.79e-202 Length: 653
 Percent Similarity: 210.00 Matches: 210
 Best Local Similarity: 100.00% Conservative: 0
 Query Match: 42.17% Mismatches: 0
 DB: 9 Indels: 0
 Gaps: 0

US-09-872-523-5 (1-498) x AU204939 (1-653)

QY 1 MetSerGluPheLeuLysIleValArgAlaAsnLysSerAspArgLysLeuAspLys 20
 DB 22 ATGTCGAATTTCTGAAAATTGTCAGACTAACAAAAATCGACAGAAAACCTCGATTAG 81
 QY 21 ThrTYrLeuTPGluSerTYrLeuHisGlnPheGluLysGlyLysThrSerPheIlePro 40
 DB 82 ACCTACTGTGGGATCCTATTATTCATCAGTTCCGAAAGGAAAACTTCTTCTTCCA 141
 QY 41 ValGluAlaPheAsnArgAsnLeuThrValAsnPheAsnGluCysValLysGluGlyVal 60
 DB 142 GTTGACGATTCACATCGTACCTTACCTTAATTTTAAACGATCGTGAAGAGAGAGT 201
 QY 61 IlePheGluThrValValHisAspTYrAspLysAsnCysAspSerIleGlnValArgTrp 80
 DB 202 ATCTTCGAACAGTGCTCATGATTAATGACAGAACTCGATTCCATTCAAGTCAGATGG 261
 QY 81 PheAlaArgIleGluLysValCysGlyTYrArgValLeuAlaGlnPheIleGlyAlaAsp 100
 DB 262 TTTGACGATTCGAAAAGTTTGGCGATACAGAGTTTGGCTCACTTATCGAGCTGAC 321
 QY 101 ThrLysPheITPLeuAsnIleLeuSerAspAspMetPheGlyLeuAlaAsnAlaMet 120
 DB 322 ACGAAATTTTGCTCAATATTTTATCGGACGATATGTTGGTTGGCAACGCCCAATG 381
 QY 121 SerAspProAsnMetAspLysIleValTYrAlaPProPoleuAlaIleAsnGluGluTYr 140
 DB 382 AGTGATTCACAAATGATGATTAATTTATATGCTCGCGCTTGCAATCAACAAGAAATTC 441
 QY 141 GlnAsnAspMetValAsnTYrValAsnAsnCysIleAspGlyGluIleValGlyGlnThr 160
 DB 442 CAATATGATATGTAATTAATTAATTAATTCATTCATTCATTCATTCATTCATTCATTC 501
 QY 161 SerLeuSerProLysPheAspGluGlyLysAlaLeuLeuSerLysHisArgPheLysVal 180
 DB 502 TCGTGCTCCCAAAATTCGATGAAGGAGCTCTCTTAACGAAACATCGTTTCAAAAGTT 561
 QY 181 GlyIlnArgLeuGluLeuLeuAsnTYrSerAsnSerThrGluIleArgValAlaArgTle 200

DB 562 GGACACACTCTTGACATTTAATTTACTCCATTTCTACTGAATACCGGTAGCCGAAAT 621
 QY 201 GlnGluIleCysGlyArgArgMetAsnVal 210
 DB 622 CAAGAAATATCTGACGACGAATGATGTA 651

RESULT 2

BI0101701 581 bp mRNA linear EST 18-JAN-2002
 LOCUS BI0101701 unpublished oligo-capped cDNA library, C. elegans L1 stage
 DEFINITION Caenorhabditis elegans cDNA clone YK1029h03 5', mRNA sequence.

ACCESSION BI0101701 GI:18244371
 KEYWORDS EST.

SOURCE
 ORGANISM

Caenorhabditis elegans
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae
 ; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE 1 (bases 1 to 581)
 Kohara, Y., Shin-I, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.

AUTHORS
 TITLE A complementary view of the C.elegans genome
 JOURNAL Unpublished
 COMMENT Contact: Tadasu Shin-I
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp.

FEATURES
 source Location/Qualifiers

1..581
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /strain="N2"
 /db_xref="taxon:6239"
 /clone="YK1029h03"
 /sex="Hermaphrodite"
 /tissue_type="whole animal"
 /dev_stage="L1"
 /clone_lib="unpublished oligo-capped cDNA library, C. elegans L1 stage"

BASE COUNT 189 a 103 c 123 g 166 t

ORIGIN

Alignment Scores:

Pred. No.: 8.65e-178 Length: 581
 Score: 186.00 Matches: 186
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 37.35% Indels: 0
 DB: 12 Gaps: 0

US-09-872-523-5 (1-498) x BI0101701 (1-581)

QY 1 MetSerGluPheLeuLysIleValArgAlaAsnLysSerAspArgLysLeuAspLys 20
 DB 22 ATGTCGAATTTCTGAAAATTGTCAGACTAACAAAAATCGACAGAAAACCTCGATTAG 82
 QY 21 ThrTYrLeuTPGluSerTYrLeuHisGlnPheGluLysGlyLysThrSerPheIlePro 40
 DB 82 ACCTACTGTGGGATCCTATTATTCATCAGTTCCGAAAGGAAAACTTCTTCTTCCA 142
 QY 41 ValGluAlaPheAsnArgAsnLeuThrValAsnPheAsnGluCysValLysGluGlyVal 60
 DB 143 GTTGACGATTCACATCGTACCTTATTCATTCATTCATTCATTCATTCATTCATTC 202
 QY 61 IlePheGluThrValValHisAspTYrAspLysAsnCysAspSerIleGlnValArgTrp 80
 DB 203 ATCTTCGAACAGTGCTCATGATTAATGACAGAACTCGATTCCATTCAAGTCAGATGG 262
 QY 81 PheAlaArgIleGluLysValCysGlyTYrArgValLeuAlaGlnPheIleGlyAlaAsp 100

```

Db      263 TTTGCAGCAATTGAAAAAGTTTGGCATACAGAGTCTCGCTCAGTTATCGAGCTGAC 322
Qy      101 ThrlyspheTrpLeuAsnIleleuSeraspMetPheGlyLeuAlaAsnAlaIleMet 120
Db      323 ACGAAATTTTGGCTCAATATTTTATTCGACGATATGTTGGTTCGCAACGCCCAATG 382
Qy      121 SerAspProAsnMetAspLysIleValTyrAlaProLeuAlaIleAsnGluGlyTyr 140
Db      383 ATGTATCCCAATATGATATAATTTGATATGCTCCGCCCTTGCAATCAACGAGAAATAC 442
Qy      141 GlnAsnAspMetValAsnTyrValAsnAsnCysIleAspGlyLeuIleValGlyGlnThr 160
Db      443 CAAATGATATGATGTAATATGTAATATGATGATGATGCGCAATTCGCGCAACT 502
Qy      161 SerLeuSerProLysPheAspGlyLysAlaLeuLeuSerLysHisArgPheLysVal 180
Db      503 TCGCTGTCTCCAAATTCGATGAAGGAGGCTCTCCTTAAGCAAGCATGTTCAAGTT 562
Qy      181 GlyGlnArgLeuGluLeu 186
Db      563 GGCACAACTCTTGAACTA 580

```

```

RESULT 3
Bj124772 557 bp mRNA linear EST 30-MAY-2003
LOCUS Bj124772 unpublished oligo-capped cDNA library, C. elegans L1 stage
DEFINITION Caenorhabditis elegans cDNA clone yk115b11 5', mRNA sequence.
ACCESSION Bj124772 GI:31246008
VERSION Bj124772.2 GI:31246008
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Eukaryota; Metazoa; Nematoda; Chromodorea; Rhabditida; Rhabditioidea
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished
COMMENT On Jan 23, 2002 this sequence version replaced gi:18284913.
Contact: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.
Location/Qualifiers
FEATURES
source
1..557
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk115b11"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, C. elegans L1 stage"

```

```

BASE COUNT 180 a 180 g 158 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 1 61e-155 Length: 557
Score: 164.00 Matches: 177
Percent Similarity: 99.44% Conservative: 0
Best Local Similarity: 99.44% Mismatches: 0
Query Match: 32.93% Indels: 1
DB: 12 Gaps: 0
US-09-872-523-5 (1-498) x Bj124772 (1-557)

```

```

Qy      2 SerGluPheLeuLysIleValArgAlaAsnLysSerAspArgLysLeuAspLysThr 21
Db      26 TCTGATTTCTGAAATTTCTCAGAGCTAACCAAAAATCGACAGAAATCTCATAGAC 85
Qy      22 TyrLeuTrpGluSerTyrLeuHisGlnPheGlyLysGly-LysThrSerPheIleProVa 41
Db      86 TACTTGGGAAATCCATTTACATCAGTTTCGAGAAAGNAAAAAATCTTTCATCTCACT 145
Qy      41 IGluAlaPheAsnArgAsnLeuThrValAsnPheAsnGluCysValLysGluGlyAlaI 61
Db      146 TGAAGCATTCATTCGTAACCTTACAGTAAATTTTAAAGAAATGCGTGAAGAGAGATAT 205
Qy      61 ePheGluThrValAlaHisAspTyrAspLysAsnCysAspSerIleGlnValArgTrpH 81
Db      206 CTTGCAAAACAGTGCATCATATATACAGAAAGTCCGATTCGATTCAGTCAGATGTT 265
Qy      81 eAlaArgIleGluLysValCysGlyTyrArgValLeuAlaGlnPheIleGlyAlaAspTh 101
Db      266 TGCACGAAATTAAGAAAGTTTGGGATACAGAGTTCGCTCAGTTTATCGAGCTGACAC 325
Qy      101 rlyspheTrpLeuAsnIleleuSeraspMetPheGlyLeuAlaAsnAlaIleMetSe 121
Db      326 GAAATTTTGGCTCAATATTTTATCGACGATATGTTGGTTCGCAACGCCCAATGAG 385
Qy      121 rAspProAsnMetAspLysIleValTyrAlaProLeuAlaIleAsnGluGlyTyrG1 141
Db      386 TGATCCCAATATGATAAATTTGATATGCTCCGCCCTTGCAATCAACGAGAAATACCA 445
Qy      141 nAsnAspMetValAsnTyrValAsnAsnCysIleAspGlyGluIleValGlyGlnThrSe 161
Db      446 AATGATATGATGAATATGTAATATGATGATGATGCGAAATGCTGCGCAAACTTC 505
Qy      161 rleuSerProLysPheAspGlyLysAlaLeuLeuSerLysHisArgPhe 178
Db      506 GCTGTCTCCAAATTCGATGAAGGAGGAGGCTCTCCTTAAGCAAGCATGCTTTC 557

```

```

RESULT 4
Bj115564 519 bp mRNA linear EST 23-JAN-2002
LOCUS Bj115564 unpublished oligo-capped cDNA library, C. elegans L1 stage
DEFINITION Caenorhabditis elegans cDNA clone yk1192n09 5', mRNA sequence.
ACCESSION Bj115564 GI:18275658
VERSION Bj115564.1 GI:18275658
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Eukaryota; Metazoa; Nematoda; Chromodorea; Rhabditida; Rhabditioidea
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished
COMMENT Contact: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.
Location/Qualifiers
FEATURES
source
1..519
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1192n09"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, C. elegans L1 stage"

```

```

BASE COUNT 170 a 170 g 109 t 149 t

```


Source

```

source
1. .761
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1029h03"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, C. elegans L1 stage"
233 a 150 c 144 g 232 t 2 others
BASE COUNT
ORIGIN

```

Alignment Scores:

Pred. No.:	5,64e-126	Length:	761
Score:	135.00	Matches:	135
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	27.11%	Indels:	0
DB:	12	Gaps:	0

US-09-872-523-5 (1-498) x BJ129361 (1-761)

QY	286	SetrAspAspValThrPheaspGlnIleuValLysaspProIleaspPrometIletrArg	305
Db	759	TCAGACGAGCTCACATTTGATCATATTACCAAAAGATCCAAATGATCCCATGATTTGGACA	700
QY	306	LysValLysValGlyGlnLysPheGlnIleuValLeaspProLeuAlaGlnIlePheasn	325
Db	699	AAAGTTAAAGGTGGACAAAAGTTTGAAGCTCATGACCCNTTGGCTCAGCAATTCATTAAC	640
QY	326	LeuHisValAlaSerIleLeuLysPheGlyLysGlyThrGlnGlyTyrLeuIleValGlyMet	345
Db	639	CTCCACGTCGGCTTCATTTCTCAAAATTTGGCAAACTGAAAGATGTCATTATGTGGGATG	580
QY	346	AspGlyProAspAlaIleuGlnLysPserPheProIleHisIleasnAsnThrPheMetPhe	365
Db	579	GATGGTCCAGATGSCACTTGAGACACGTTTTCTCATTCATCAATTAATCAATTAATGATTC	520
QY	366	ProValGlyTyrAlaGlnLysLysTyrAsnLeuGlnIleuValProPheAspGluPheLysGly	385
Db	519	CCAGTTGGTATTGCGGAAAGATGATTAATTTGGAACTGTGTCCGCGAGATGAGTTCAAAGCA	460
QY	386	ThrPheAspGlyTyrAspGluTyrLeuGlnLysGlnSerAlaGlnThrIleuProLeuAspLeu	405
Db	459	ACATTCAGATGGAGATGAACTACTTGAGGAAAGAAATCTGAGAAACCTTACCGCTTGACTTG	400
QY	406	PheLysProMetProSerGlnGlnAlaGlnLysPheLysVal	420
Db	399	TTTCAGCCCAATGCCCTTCCCAAGAGGATTTAACAATTTTAAAGTA	355

RESULT 7

[illegible]JOURNAL
COMMENT

JOURNAL
COMMENT
Nat. Genet., (2003) In press
Contact: Vidal M
Marc Vidal Laboratory
Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739
Email: Marc_vidale@fci.harvard.edu
Sequence tag of Gateway entry clones. The primers used were
designed on the predicted protein encoding ORF. C. elegans ORFome
cloning project : Contact david_hille@fci.harvard.edu or
marc_vidale@fci.harvard.edu
POLYA=NO.

FEATURES	Location/Qualifiers
source	1. .607

```

/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/sex="Hermaphrodite and male"
/tissue_type="whole animal"
/dev_stage="mixed stage"
/clone_idb="AD-wrmCDNA"

```

RNA isolated from both hermaphrodite and male N2 worms of all larval stages, embryos, adults and dauvers and the subsequent generation of CDNAs by poly(A) priming. The CDNAs were cloned into pPC86"

Alignment Scores:

Pred. No.:	5,33e-119	Length:	607
Score:	128.00	Matches:	120
Percent Similarity:	100.00%	Conservative:	0
Best local Similarity:	100.00%	Mismatches:	0
Query Match:	25.70%	Indels:	0
DB:	14	Gaps:	0

US-09-872-523-5 (1-498) x CB404801 (1-607)

QY	29	GINLEALALYASAPROILAEAPROMETILERPARGLYSVALYLSVALGILGILNLS	312
Db	573	CAATTTAGCAAAAGATCCAAATGATCCCATGATTTGGAGAAAAGTTAAGGTGGACAAAAG	514
QY	313	PhEGLIEULAEAPROLEULAGLINCINPHEASNASINLEUHSIVALALASERILEU	332
Db	513	TTTGAGCTATGAGACCCCTTGGCTGACGAATTCATTAACCTCCAGTCGCTTGCAATTC	454
QY	333	LysPHECysLysThrGLINGLTYRLEULLEVALGILELSPGLIPIROSPALALEUL	352
Db	453	AAATTTTGCAGAAACTGAAAGATATCTTATTTGCGAAGATGATGGTCCAGATGCACCTGAA	394
QY	353	ASPSErPHEPROILIEHLSLEASNSNRHPHEMETPHEPROVALGLYTYRILAGILNLS	372
Db	393	GACAGTTTTCATATCATATCATATATATACATTATATGTTCCAGTTGGTTATGGCGAAAAG	334
QY	373	TYrASUleUGLIEULVALPROBPASPLUPHELYSGLYTHRPhEATGTPASPLUTYr	392
Db	333	TATAAATTTGGAACTGTTCGCCCAATAGATTCCAAAGAACATTCAATGGGATGAAATAC	274
QY	393	LeUGILYSGLSERALAGLUTHrLEUDROLEUNASPLEUPHELYSPROMETPROSErGIN	412
Db	273	TTTGAGAAAGATCTGACAGAAACCTTACCGCTTGACTGTTCAGACCAATGCCITGCCAA	214
QY	413	GLUARGLEUASPLYSPHELYSVAL	420
Db	213	GACGAGTTTAGCAAAATTTAAAGTA	190

ATTN:

TITLE
C. elegans Unrome Version 1.1: experimental verification of
genome annotation and resource for proteome-scale protein
expression

Caenorhabditis elegans cDNA clone yk292d9 5', mRNA sequence.
ACCESSION C42305
VERSION C42305.1 GI:2378542
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
AUTHORS Kohara,Y., Motolashi,T., Tabara,H., Watanabe,H., Sugimoto,A., Sano
M., Miyata,A. and Nishigaki,A.
TITLE Expression map of the C.elegans genome
JOURNAL Unpublished
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers
FEATURES
source 1..360
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk292d9"
/sex="hermaphrodite"
/dev_stage="embryo"
/clone.lib="Yuji Kohara unpublished cDNA:Strain N2
hermaphrodite embryo"
BASE COUNT 119 a 62 c 78 g 101 t
ORIGIN
Alignment Scores:
Pred. No.: 3,76e-110 Length: 360
Score: 119.00 Matches: 119
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 23.90% Indels: 0
DB: 14 Gaps: 0
US-09-872-523-5 (1-498) x C42305 (1-360)
QY 9 ArgAlaasnLysLysSerAspArgLysLeuaspLysThrTyrLeuTrpGluSerTyrLeu 28
DB 2 AGAGCTAACCAAAATCGACAGAAACTCGATAAGCTACTGTGGAAATCGATTTTA 61
QY 29 HisGlnPheGluLysGlyLysThrSerPheIleProValGluAlaPheAsnArgAsnLeu 48
DB 62 CATCAGTTCGAAAGGAAAACTCTTTCATTCAGTTGAAGCATTCATCGTAACCTT 121
QY 49 ThrValAsnPheAsnGluCysValLysGluGlyValIlePheGluThrValAlaHisasp 68
DB 122 ACAGTTAATTTTAACGAATGCGTGAAGAGAGATTATCTCGAAACAGTGTCCATGAT 181
QY 69 TyrAspLysAsnGlySerPheIleGlnValArgTppPheAlaArgIleGluLysValCys 88
DB 182 TATGACAAGAACTGCGATTGCAATGACAGATGCTTGGACGAATGAAAAAGTTTGC 241
QY 89 GilyTyrArgValLeuAlaGlnPheIleGlyAlaAspThrLysPheTrpLeuAsnIleLeu 108
DB 242 GGATACAGAGTTCTGGCTCAGTTTATCGAGCTGCACAGAAATTTTGGCTCAATTTTAA 301
QY 109 SerAspAspMetPheGlyLeuAlaAsnAlaAlaMetSerAspProAsnMetAspLys 127
DB 302 TCGACGAGATTTGTTGGTTCGAACGCGCAATGAGATGCCCAATATGATATAA 358
RESULT 9 AV197399 380 bp mRNA linear EST 26-JUL-1999
LOCUS AV197399 Yuji Kohara unpublished cDNA:Strain N2 hermaphrodite
DEFINITION embryo Caenorhabditis elegans cDNA clone yk65h11 5', mRNA

sequence.
ACCESSION AV197399
VERSION AV197399.1 GI:5581170
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
AUTHORS Kohara,Y., Shin,I.,T., Thierry-Mieg,J., Thierry-Mieg,D., Mitsuki,H.,
Nishigaki,A., Motolashi,T., Zeng,Q., Watanabe,H., Sugimoto,A., Sano
M., Miyata,A., Mitani,Y., Ida,K., Uesugi,H., Sugiyama,Y. and
Nomoto,H.
TITLE Expressed genes in C.elegans
JOURNAL Unpublished
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers
FEATURES
source 1..380
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk65h11"
/sex="hermaphrodite"
/dev_stage="embryo"
/clone.lib="Yuji Kohara unpublished cDNA:Strain N2
hermaphrodite embryo"
BASE COUNT 125 a 63 c 83 g 108 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 4,96e-103 Length: 380
Score: 112.00 Matches: 125
Percent Similarity: 99.21% Conservative: 0
Best Local Similarity: 99.21% Mismatches: 0
Query Match: 22.49% Indels: 1
DB: 9 Gaps: 0
US-09-872-523-5 (1-498) x AV197399 (1-380)
QY 6 LysIleValAlaGalaAsnLysSerAspArgLysLeuaspLysThrTyrLeuTrp-GI 25
DB 3 AAAATTGTCAGAGCTAACCAAAATCGACAGAAACTCGATAAGACCTACTGTGGGA 62
QY 25 uSerTyrLeuHisGlnPheGluLysGlyLysThrSerPheIleProValGluAlaPheAs 45
DB 63 ATCTATTATTACATCGCTTCGAGAAAGAAAACTCTTTCATTCAGTTGAAGCATTCGA 122
QY 45 nArgAsnLeuThrValAsnPheAsnGluCysValLysGluGlyValIlePheGluThrVa 65
DB 123 TCGTAACCTTACAGTTAATTTTAACGAATGCGTGAAGAGAGATTATCTCGAAACAGT 182
QY 65 lValHisaspLysAspLysAsnGlySerPheIleGlnValArgTppPheAlaArgIleG1 85
DB 183 GGTCCATGATTAATGACAAGAACTGCGATTGCAATGACCAATGCTTGGACGAATTGA 242
QY 85 uLysValCysGlyTyrArgValLeuAlaGlnPheIleGlyAlaAspThrLysPheTrpLe 105
DB 243 AAAAGTTTGGGATACAGAGTTCTGGCTCAGTTTATCGAGCTGCACAGAAATTTTGGCT 302
QY 105 uAsnIleLeuSerAspAspMetPheGlyLeuAlaAsnAlaAlaMetSerAspProAsnMe 125
DB 303 CAATATTTTATCGACGAGATATGTTGTTGGCAACGCGCAATGAGATGCCCAATAT 362
QY 125 tAspLysIleValTyr 130
DB 363 GGATAAAATTTGTATAT 378

[illegible]

RESULT 11	BJ153087	686 bp	mRNA	linear	EST 24-JAN-2002
LOCUS	BJ153087/c				
DEFINITION	BJ153087 unpublished oligo-capped cDNA library; C. elegans L1 stage				
ACCESSION	Caeonrhabditis elegans cdna clone yk1315b11 3'				
VERSION	BJ153087				
KEYWORDS	BJ153087.1 GI:18321072				
SOURCE	EST.				
ORGANISM	Caeonrhabditis elegans Caeonrhabditis elegans Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea				
REFERENCE	1 (bases 1 to 686)				
AUTHORS	Kohara,Y., Shn-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y. and Sugano,S.				
TITLE	A complementary view of the C.elegans genome				
JOURNAL	Unpublished				
COMMENT	Contact: Tadenu Shin-i Center For Genetic Resource Information National Institute Of Genetics 111 Yata, Mishima, Shizuoka 411-8540, Japan Tel.: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp.				
FEATURES	Location/Qualifiers				
source	1..686 /organism="Caenorhabditis elegans" /mol_type="mRNA" /strain="N2" /db_xref="taxon:6239" /clone="yk1315b11" /sex="hermaphrodite" /tissue_type="whole animal" /dev_stage="L1" /clone_idb="unpublished oligo-capped cdna library, C. elegans L1 stage"				
BASE COUNT	214 a 133 c 131 g 204 t				4 others
ORIGIN					
Alignment Scores:					
Pred. No.:	1.13e-98				Length: 686
Score:	108.00				Matches: 108
Percent Similarity:	100.00%				Conservative: 0
Best Local Similarity:	100.00%				Mismatches: 0
Query Match:	21.69%				Indels: 0
DB:	12				Gaps: 0
US-09-872-523-5 (1-498) x BJ153087 (1-686)					
OY	313 PheglutleuIleaapProleuaAGlnInPheasnsAenleuHIsvalAlaSerIleu				332
Db	684 TTGGGCGCATGCAGCCCTTGTGGCTTCAGCAATTAACCTCCACGTCGCTTGCTGC				625
OY	333 LysPheCyLysTThGlUGlTYTrLeuIleValGIymeTaSPglProSPAalaLeuGLu				352
Db	624 AAATTTTCCAACAACGAAGAATATCTTTGTGGGAAGAATGATGCCAGATGCACCTTAA				565
OY	333 AspserPheProlleHisIleasnAsnThrPhemetPheProvalGIlyTyralaGIuLys				372
Db	564 GACAGTTTTCCTATTCATATCAATAATATCAATTATGTGCCAGGTTGATGCGGAAGAAG				505
OY	373 TyraSenleugluIleValProProaspIuPhelysgLYrrPheArgTTPrasgLUyrr				392
Db	504 TATAATTTGGAACTGTGTCCCGCCAGATGAGTTCAAAGAACATTCAGATGGATGAATAC				445
OY	393 LeugluLysgluSerAlaaglUthrLeuProLeuaspleuPhelysPrometProSeGln				412
Db	444 TTGGAGAAAGAATTCGCGAANAACCTTACCCTTGACTTGTTCAGAGCAATGCTTCCCA				385
OY	413 GluarGluasplysPheLysval 420				
Db	384 GAGAGATTAGACAATTTTAAGGTA 361				

RESULT 12
AUI11588 340 bp mRNA linear EST 19-OCT-2000
LOCUS AUI11588
DEFINITION AUI11588 unpublished oligo-capped cDNA library Caenorhabditis
elegans cDNA clone yk7325 5', mRNA sequence.
ACCESSION AUI11588
VERSION AUI11588.1 GI:10925155
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 340)
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C. elegans genome
JOURNAL Unpublished
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.

FEATURES
Source Location/Qualifiers
1..340
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk7325"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="varied"
/clone_11b="unpublished oligo-capped cDNA library"

BASE COUNT 111 a 58 c 72 g 98 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 5.88e-93 Length: 340
Score: 102.00 Matches: 102
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.48% Indels: 0
DB: Gaps: 9

US-09-872-523-5 (1-498) x AUI11588 (1-340)

QY 2 SerGIuPheLeuLysIleValArgAlaAsnLysLysSerAspArgLysLeuAspLysThr 21
DB 34 TCTGAATTTCTGAAATATGTCAGACCTAACAAATAATCGACAGAAACTCGATAGACC 93
QY 22 TyrLeuTPGiuserTyrLeuHisGlnPheGluLysGlyLysThrSerPheIleProVal 41
DB 94 TACTGTGGAACTCTATTATTCATCATGTCGAGAAAGAAACTTCTTCAATTCACAGTT 153
QY 42 GluAlaPheAsnArgAsnLeuThrValAsnPheAsnGluLysValLysGluLysIle 61
DB 154 GAGACATTCACATCGTAACCTTACACTTAATTTAAGCAATGCGTGAAGAAAGACATTATC 213
QY 62 PheGluThrValValHisAspTyrAspLysAsnGlySerSerIleGlnValArgTrpPhe 81
DB 214 TTCGAACAGTGGTCATCATGATATGACAAAGACTCGATTGATTCACAGTCAAGTGGTTT 273
QY 82 AlaArgIleGluLysValCysGlyTyrArgValIleuAlaGlnPheIleGlyAlaAspThr 101
DB 274 GCACCAATTTGAATAAGTTTGGCGATACAGATTCTGCTCAATTTATCGACCTACACG 333
QY 102 LysPhe 103
DB 334 AAATTT 339

RESULT 13
AUI16296/c 300 bp mRNA linear EST 19-OCT-2000
LOCUS AUI16296
DEFINITION AUI16296 unpublished oligo-capped cDNA library Caenorhabditis
elegans cDNA clone yk743a4 3', mRNA sequence.
ACCESSION AUI16296
VERSION AUI16296.1 GI:10929863
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 300)
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C. elegans genome
JOURNAL Unpublished
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.

FEATURES
Source Location/Qualifiers
1..300
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk743a4"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="varied"
/clone_11b="unpublished oligo-capped cDNA library"

BASE COUNT 95 a 63 c 54 g 88 t
ORIGIN

Alignment Scores:
Pred. No.: 1.26e-62 Length: 300
Score: 72.00 Matches: 72
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.46% Indels: 0
DB: Gaps: 9

US-09-872-523-5 (1-498) x AUI16296 (1-300)

QY 427 ValGIuLeuArgLeuGluAlaAlaAspMetCysGluAsnGlnPheIleCysProAlaThr 446
DB 299 GTAGAGCTACGCGCTTGAAAGCTGCTGACATGTGTGAATAATCAGTTATTGTCACCTACA 240
QY 447 ValLysSerValHisGlyArgLeuIleAsnValAsnPheAspGlyTrpAspGluLysPhe 466
DB 239 GTGAATCAGTTCAAGCAAGACTGATTAATGATTCATTCGACGCTGGGATGACAAATTT 180
QY 467 AspGluLeuTyrAspValAspSerHisAspIleLeuProIleGlyTrpCysGluAlaHis 486
DB 179 GATGAACGTGATGATGATGAGTCCATCATATATTCACGATGATGATGATGATGATGATGAT 120
QY 487 SerTyrValLeuGlnProProLysLysTyrAsnTyr 498
DB 119 AGTTATGTTCTACCACTCCGAAAGATGACACTAT 84

RESULT 14
AUI15428/c 300 bp mRNA linear EST 19-OCT-2000
LOCUS AUI15428
DEFINITION AUI15428 unpublished oligo-capped cDNA library Caenorhabditis
elegans cDNA clone yk7325 3', mRNA sequence.
ACCESSION AUI15428
VERSION AUI15428.1 GI:10928995
KEYWORDS EST.
SOURCE Caenorhabditis elegans

ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
; Rhabditidae; Pelodderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 300)
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C. elegans genome
JOURNAL Unpublished
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers
1. 300
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk732e5"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="varied"
/clone_lib="unpublished oligo-capped cDNA library"
BASE COUNT 94 a 60 c 54 g 89 t 3 others
ORIGIN
Alignment Scores:
Pred. No.: 1,34e-60 Length: 300
Score: 70.00 Matches: 70
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.06% Indels: 0
DB: 9 Gaps: 0
US-09-872-523-5 (1-498) x AU115428 (1-300)
QY 429 LeuAtgLeuGluAlaAlaAspMetCysGluAsnGlnPheIleCysProAlaThrValys 448
|||||
DB 298 CTACGCCCTGAGAGCTGCTGACATGTGTGAATAATCATTTATTGTCTCAGCTACAGGANA 239
QY 449 SerValHisGlyArgLeuIleAsnValAsnPheAspGlyTrpAspGluGluPheAspGlu 468
|||||
DB 238 TCAGTTCATGAGAGAGATGATTAATGTCATTTGACGCTGGGATGAGAAATTTGATGAA 179
QY 469 LeuTyAspValAspSerHisAspIleLeuProIleGlyTrpCysGluAlaHisSerTy 488
|||||
DB 178 CTGTATGATGTGACTCCCATGATATTTTACCGATAGAGATGTGTGAGCGCACAGTTAT 119
QY 489 ValLeuGlnProProlysIleTyAsnTy 498
|||||
DB 118 GTTCTACACCTCCGAAAAAGTACAACTAT 89
RESULT 15
AU112450 337 bp mRNA linear EST 30-MAY-2003
LOCUS AU112450
DEFINITION AU112450 unpublished oligo-capped cDNA library Caenorhabditis
elegans cDNA clone yk743a4 5', mRNA sequence.
ACCESSION AU112450
VERSION AU112450.2 GI:31237498
KEYWORDS EST.
SOURCE
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
; Rhabditidae; Pelodderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 337)
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C. elegans genome
JOURNAL Unpublished
COMMENT On Oct 19, 2000 this sequence version replaced gi:10926017.

Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers
1. 337
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk743a4"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="varied"
/clone_lib="unpublished oligo-capped cDNA library"
BASE COUNT 107 a 59 c 74 g 92 t 5 others
ORIGIN
Alignment Scores:
Pred. No.: 1.97e-52 Length: 337
Score: 62.00 Matches: 62
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.45% Indels: 0
DB: 9 Gaps: 0
US-09-872-523-5 (1-498) x AU112450 (1-337)
QY 1 MetSerGluPheLeuIleValArgAlaAsnIleCysSerAspArgIleAspIys 20
|||||
DB 23 ATGCTGGAATTTCTGAAAATTTGTCAGAGCTAACCAAAAAATCGACAGAAAATCGATTAAG 82
QY 21 ThrTyLeuTrpGluSerTyIleuHisGlnPheGluGlyIleThrSerPheIlePro 40
|||||
DB 83 ACCTACTTGTGGAAATCTATTTCATCAGTTTCGAGAAAGAAAACCTCTTCATTTCA 142
QY 41 ValGluAlaPheAsnArgAsnLeuThrValAsnPheAsnGluCysValIleGluGlyVal 60
|||||
DB 143 GTTGAAGCAATTCATGCTAACCTTACAGTTAATTTTAACGAATGCCGTGAAGAGAGATT 202
QY 61 IlePhe 62
|||||
DB 203 ATCTTC 208

Search completed: September 2, 2003, 23:10:26
Job time : 3076 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: September 2, 2003, 17:58:40 ; Search time 407 Seconds
(without alignments)
3302.998 Million cell updates/sec

Title: US-09-872-523-5

Perfect score: 498
Sequence: 1 MSEFLKIVRANKSDRKLDK.....PICGCAHSYVLQPEKKYNY 498

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 252756 seqs, 1349719017 residues

Word size: 1

Total number of hits satisfying chosen parameters: 5103490

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODE=frame-p2n.model -DEV=x1h
-O/cgcn2.1/USPTO.spool/US09872523/runat.02092003.073000.27141/app.query.fasta.1.647
-DB=N_Geneseq_19Jun03_QEWT-fastap_SUFFIX=coll.ring_WITHMATCH=0.1 -LOOFCLE-
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09872523.ecgn.1.1.312.0runat.02092003.073000.27141 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLCK=100 -LONGLOC
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : N_Geneseq_19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	498	100.0	1497	24	ABL52497	Caenorhabditis ele
2	411	82.5	1497	24	ABL52515	C. elegans mutant
3	411	82.5	1497	24	ABL52516	C. elegans mutant
4	397	79.7	1497	24	ABL52517	C. elegans mutant
5	278	55.8	2307	24	ABL52518	C. elegans lin-61
6	278	55.8	2307	24	ABL52519	C. elegans lin-61
7	278	55.8	2307	24	ABL52520	C. elegans lin-61
8	9	1.8	344	21	AAA31397	Plant microsatelli
9	9	1.8	350	21	AAA31396	Plant microsatelli
10	9	1.8	351	21	AAA31294	Plant microsatelli
11	9	1.8	363	21	AAA31381	Plant microsatelli
12	9	1.8	367	21	AAA31392	Plant microsatelli
13	9	1.8	383	21	AAA31276	Plant microsatelli
14	9	1.8	401	21	AAA31391	Plant microsatelli
15	9	1.8	407	21	AAA31390	Plant microsatelli
16	9	1.8	410	21	AAA31259	Plant microsatelli
17	9	1.8	412	21	AAA31387	Plant microsatelli
18	9	1.8	413	21	AAA31410	Plant microsatelli
19	9	1.8	423	21	AAA31314	Plant microsatelli
20	9	1.8	425	21	AAA31326	Plant microsatelli
21	9	1.8	426	21	AAA31403	Plant microsatelli
22	9	1.8	430	21	AAA31336	Plant microsatelli
23	9	1.8	433	21	AAA31284	Plant microsatelli
24	9	1.8	450	21	AAA31250	Plant microsatelli
25	9	1.8	457	21	AAA31269	Plant microsatelli
26	9	1.8	479	21	AAA31291	Plant microsatelli
27	9	1.8	479	21	AAA31292	Plant microsatelli
28	9	1.8	481	21	AAA31306	Plant microsatelli
29	9	1.8	484	21	AAA31305	Plant microsatelli
30	9	1.8	502	21	AAA31309	Plant microsatelli
31	8	1.6	234	25	ABX31222	Human GDP-mannose
32	8	1.6	843	23	AA571335	DNA encoding novel
33	8	1.6	872	22	AAH04590	Human cDNA clone (
34	8	1.6	1250	24	ABO41056	Oligonucleotide fo
35	8	1.6	1250	24	ABO41057	Oligonucleotide fo
36	8	1.6	1442	22	AAH37421	Human secreted pro
37	8	1.6	2202	22	AAH17401	Human cDNA sequenc
38	8	1.6	2271	15	AAO62176	Human TLE-2 gene.
39	8	1.6	2271	15	AAO62176	Human TLE-2 gene.
40	8	1.6	2272	23	AA568845	DNA encoding novel
41	8	1.6	2537	20	AAV80583	Kidney injury asso
42	8	1.6	2719	22	AAH13886	Human cDNA sequenc
43	8	1.6	2943	24	ABN85369	Sputumretrovirus DN
44	8	1.6	3524	20	AAH80633	Stromal cell-deriv
45	8	1.6	3524	20	AAH76381	Human stromal cell

ALIGNMENTS

RESULT 1

ABL52497 standard; DNA; 1497 BP.

ABL52497; (first entry)

17-JUL-2002

Caenorhabditis elegans lin-61 nucleotide sequence SEQ ID NO:6.

Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;

cell proliferation; nematode; cancer; gene; ds.

Caenorhabditis elegans.

Key Location/Qualifiers
CDS 1..1497

```

FT      /*tag= a
FT      /product= "LIN-61 protein"
XX      MO200194545-A2.
XX      13-DEC-2001.
XX      01-JUN-2001; 2001MO-US17909.
XX      02-JUN-2000; 2000US-208802P.
XX      (MAST ) MASSACHUSETTS INST TECHNOLOGY.
PI      Horvitz HR, Davison EM, Lu X;
XX      WPI: 2002-401590/43.
DR      P-PSDB; ABB78650.
XX      New Caenorhabditis elegans nucleic acid involved in tumor suppressor
PT      pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
PT      diagnosing and treating cell proliferative diseases such as cancer
XX      Claim 24; Page 74; 116pp; English.
XX      The present invention describes a substantially pure nematode C. elegans
CC      (Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide
CC      (see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide
CC      (see ABB78650), where the polypeptides comprise at least 130, 110, 130
CC      contiguous amino acids of the 386, 322 and 498 amino acid sequences given
CC      in ABB78648 to ABB78650, and modulate cell proliferation. (I) has
CC      cytoskeletal activity, and can be used in gene therapy. (I) is useful for
CC      modulating proliferation of a cell, and for identifying the compound that
CC      modulates cell proliferation. (I) can be used for diagnosing an animal
CC      (preferably, human) for the presence of the cell proliferation disease,
CC      or an increased chance of developing the disease, by measuring LIN-8,
CC      LIN-56 or LIN-61 nucleic acid expression in a sample obtained from the
CC      animal, by measuring LIN-8, LIN-56 or LIN-61 polypeptide in the sample.
CC      using Ab or measuring amount of LIN-8, LIN-56, LIN-61 mRNA in the sample.
CC      (I) is also useful for identifying the compound that modulates cell
CC      proliferation. (I) and its mutants are useful for understanding cell
CC      proliferative diseases including cancer, as well as in diagnosing and
CC      treating cell proliferative diseases. (I) is also useful for identifying
CC      tumor suppressors in other species such as mammals and may be used to
CC      identify therapeutic compounds. The present sequence encodes C. elegans
CC      LIN-61 from the present invention.
XX      Sequence 1497 BP; 497 A; 272 C; 312 G; 416 T; 0 other:
SQ
Alignment Scores:
Pred. No.: 0 Length: 1497
Score: 498.00 Matches: 498
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0
US-09-872-523-5 (1-498) x ABL52497 (1-1497)
OY      1 MetSerGluPheLeuAlaValArgAlaAsnLysSerAspArgLysLeuAspLys 20
DB      1 ATGCTGAAATTTCTGAAAATTTGTCAGAGCTAACAAAAATCGCAGCAAAAACTCGATAG 60
OY      21 ThrTyrLeuTrpGluSerTyrLeuHisGlnPheGluLysGlyLysThrSerPheIlePro 40
DB      61 ACCGACTGTGGGAAATCCTATTACATCAGTTTCGAGAAAGAAAACCTTTTCATTTCA 120
OY      41 ValGluAlaPheAsnArgAsnLeuThrValAsnPheAsnGluCysValLysGluGlyVal 60
DB      121 GTTGAAGCATTCATCGTAACCTTACAGTTAATTTTAAAGAAATGCGTGAAGAGAGACTT 180
OY      61 IlePheGluThrValIleValHisAspTyrAspLysAsnCysAspSerIleGluValArgTrp 80
DB      181 ATCTTCGAACAGTGTCCATGATTATGACAAAGACTCGATTCGATTTCAAGTCAGATGG 240

```

```

OY      81 PheAlaArgIleGluLysValCysGlyTyrArgValLeuAlaGlnPheIleGlyAlaAsp 100
DB      241 TTTCACAGAAATGAAAAAGTTTCGCGATACAGAGTTTCGCTCAGTTATTCGAGAGTCAC 300
OY      101 ThrLysPheTrpLeuAsnIleLeuSerAspAspMetPheGlyLeuAlaAsnAlaIleMet 120
DB      301 ACGAAATTTTGGCTCAATATTTATTCGAGAGATATGTTGGTGGCAAAAGCCCGCAATG 360
OY      121 SerAspProAsnMetAspLysIleValTyrAlaProProLeuAlaIleAsnGluTyr 140
DB      361 AGTATCCCAATATGATTAATAATTTGATATGTCGCCGCTTGCAATCAACGAAGATAC 420
OY      141 GlnAsnAspMetValAsnTyrValAsnAsnCysIleAspGlyGluIleValIleGlnThr 160
DB      421 CAAATGATATGTGTAATTTATGTAATAATTTGATGATGATGCGCAAAATCGTCGGCCAAACT 480
OY      161 SerLeuSerProLysPheAspGluGlyLysAlaLeuLeuSerLysHisArgPheLysVal 180
DB      481 TCGCTGTCTCCAAAATTCGATGAAAGGAGAGCTCTCTTAAGCAAGATCGTTCCAAAGTT 540
OY      181 GlyGlnArgLeuGluLeuLeuAsnTyrSerAsnSerThrGluIleArgValAlaArgIle 200
DB      541 GGACAAAGCTGTGAACTATTAAATTTCCAAATTCCTACTGAAATACCGTACGCGCAAT 600
OY      201 GlnGluIleCysGlyArgArgMetAsnValSerIleThrLysLysAspPheProGluSer 220
DB      601 CAAGAAATATGTGGACGACAGCAAAATGATCTATCAAAAGAAAGACTTTCCGAAATCG 660
OY      221 LeuProAspAlaAspAspArgGluValAlaPheSerSerGlySerGlnTyrTrpIleAsp 240
DB      661 CTTCACATGACAGATGACGACAGACAGCAAGCTTTAGCTCGATGATCAATTTGGATAGAC 720
OY      241 GlnGlySerPhePheIlePheProValIleGlyPheAlaValAsnGlyTyrGlnLeuAsn 260
DB      721 GAGGAGAGCTTCTCATATTTCCGTGGATTTGCACAGACAGTCATGATCAACTAAAT 780
OY      261 AlaLysGlyGluTyrIleGlnHisThrAsnLysIleGlnAlaIleLysAsnGlyGlu 280
DB      781 GCGAAAAAGGAATATATGATGACACACAAATTAATCTCAAGCAAAATTAATGAGAA 840
OY      281 AsnProArgTyrAspSerAspAspValThrPheAspGlnLeuAlaLysAspProIleAsp 300
DB      841 AATCCAAATATGATGACTCAGACGACGATTCATTCATTAATGCAAAAGATCCAAATGAT 900
OY      301 PrometIleTrpArgLysValLysValIleGlyLysPheGluLeuIleAspProLeuAla 320
DB      901 CCCATGATTTGGAGAAAAAGTTAAAGTTGGACAAAGTTTGAGCTCATCGACCCCTGGCT 960
OY      321 GlnGlnPheAsnAsnLeuHisValAlaSerIleLeuLysPheCysLysThrGluGlyTyr 340
DB      961 CAGCAATTCATTAACCTCCACGCTGCTTCGATTTCCAAATTTTGCAAAACCTGAAGATAT 1020
OY      341 LeuIleValGlyMetAspGlyProAspAlaLeuGluLysSerPheProIleHisIleAsn 360
DB      1021 CTATTTGGGGAATGAGATGCTCCAGATGCAATGCAAGCAAGTTTCCATTCATATCAAT 1080
OY      361 AsnThrPheMetPheProValIleGlyTyrAlaGluLysTyrAsnLeuGluLeuValProPro 380
DB      1081 AATACATTTATATGTTCCAGATGTTATGCCGAAAGATATATTTGGAACCTGTTCGCCGA 1140
OY      381 AspGluPheLysGlyThrPheArgTyrAspGluLysTyrLeuGluLysGluSerIleGluThr 400
DB      1141 GATAGAGTTCAAAAGAAACATTCAGATGAGTGAATTAATCTGGAGAAAGAAATCGCGAAACC 1200
OY      401 LeuProLeuAspLeuPheLysProMetProSerGlnGluArgLeuAspLysPheLysVal 420
DB      1201 CTACCGCTGACTTGTCTTCAACCCATGCTCTTCCAAAGAGATTAAGCAAAATTTAAAGTA 1260
OY      421 IleLeuIleSerLysArgValIleGlyLeuArgLeuAlaAlaAspMetCysGluAsnGln 440
DB      1261 ATTCTGATTTCCAAACGGGTAGGAGACTACGCTTGAAGCTGCTGACATGTGTGAATAATCAG 1320

```

QY	441	PhelleyCysProAlaIaThnVallySSerValHisGlyArgLeuIleAsnValAsnPheAsp	460
Db	1321	TTTTATTGTCACAGTACAGTGAATCACTTTCATGGAAGACGATAAATGTCATTAATTGCAC	1380
QY	461	GlyTrpAspGluGluPheAspGluLeuTyrAspValAspSerHisAspIleLeuProIle	480
Db	1381	GGCTGGCATGAAGAATTGATGCAACTGTATGATGTGCACTCCATGATATTTCTACCCGATA	1440
QY	481	GlyTrpCysGluAlaHisSerTyrValLeuGlnProProlYsTyrAsnTyr	498
Db	1441	GGATGGTGTGAAGCGCACAGCTTATGTTCTACAACTCCGAAAGATCAACTAT	1494
RESULT 2			
ABL52515			
ID	ABL52515	standard; DNA; 1497 BP.	
XX	ABL52515;		
AC	17-JUL-2002	(first entry)	
DT			
XX	C. elegans mutant lin-61 DNA sequence lin-61(n3446)	SEQ ID NO:73.	
DE			
XX	Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;		
KW	cell proliferation; nematode; cancer; mutant; gene; ds.		
KW			
XX	Caenorhabditis elegans.		
OS			
XX	Key	Location/Qualifiers	
FF	CDS	1..1497	
FT		/*tag= a	
FT		/product= "LIN-61(n3446) protein"	
FT			
XX	WO200194545-A2.		
PN			
XX	13-DEC-2001.		
PD			
XX	01-JUN-2001; 2001WO-US17909.		
PF			
XX	02-JUN-2000; 2000US-208802P.		
PR			
XX	(MASI) MASSACHUSETTS INST TECHNOLOGY.		
PA			
XX	Horvitz HR, Davison EM, Lu X;		
PL			
XX	WPI: 2002-401590/43.		
DR	P-PSDB; ABB78697.		
DR			
XX			
PT	New Caenorhabditis elegans nucleic acid involved in tumor suppressor		
PT	pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for		
PT	diagnosing and treating cell proliferative diseases such as cancer		
XX			
PS	Claim 27; Page 112-113; 116pp; English.		

CC	C. elegans mutant LIN-61 protein from the present invention.
xx	
sq	Sequence 1497 BP; 497 A; 271 C; 312 G; 417 T; 0 other;
Alignment Scores:	
Pred. No.:	0
Score:	411.00
Percent Similarity:	99.60%
Best Local Similarity:	99.60%
Query Match:	82.53%
DB:	24
	Gaps: 0
US-09-872-523-5 (1-498) x ABUS52515 (1-1497)	
QY	1 MetSerGluPheLeuLysIleValArgAlaAsnLysLysSeraspArgLysLeuAspLys 20
DB	1 ATGCTCGAATTTCTGAAATTTGTCAAGAGCTAACCAAAAATCGACAGCAAAACTCGATTAC 60
QY	21 ThrTYrLeuTrpGluSerTYrLeuHisGlnPheGluLysGlyLysThrSerPheIlePro 40
DB	61 ACCCTACTGTGGGAATCCTATTATTCATCATCGATTCGAGAAAGAAAACCTCTTCATTCCA 120
QY	41 ValGluAlaPheAsnArgAsnLeuThrValAsnPheAsnGluCysValLysGluGlyVal 60
DB	121 GTTGAAGCAATTCAAATCGTACCTTACAGTTAAATTTTAAAGAAACCGTGAAGAGAGATT 180
QY	61 IlePheGluThrValValHisAspTYrAspLysAsnCysAspSerIleGlnValArgTrp 80
DB	181 ATCTTCGAACAGCTGTGCTCATGATTATGACAAAGACTGGCATTCGATTCAAGTCAATGG 240
QY	81 PheAlaArgIleGluLysValCysGlyTYrArgValLeuAlaGlnPheIleGlyAlaAsp 100
DB	241 TTTGCACGAATTAAGAAAAGTTGGCGATACAGATCTGCGCTACGATTATCGGAGCTGAC 300
QY	101 ThrLysPheTrpLeuAsnIleLeuSerAspAspMetPheGlyLeuAlaAsnAlaIleMet 120
DB	301 ACGAAATTTTGGCTCAATATTTTATTCGCGAGATGTGTTGGTGGCAAAAGCCGCAATG 360
QY	121 SerAspProAsnMetAspLysIleValTYrAlaProProLeuAlaIleAsnGluGluTYr 140
DB	361 ACTGATCCCAATATGATTAATTTGATATGCTCCGCCCTGCCAATCAACGAAGAATAC 420
QY	141 GlnAsnAspMetValAsnTYrValAsnAsnCysIleAspGlyGluIleValGlyGlnThr 160
DB	421 CAAATGATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 480
QY	161 SerLeuSerProLysPheAspGluGlyLysAlaLeuLeuSerLysHisArgPheLysVal 180
DB	481 TCGCTGTCTCCAAATTTGCATAGAGGAAGGCTCTCTTAAGCAAGCATGTTTCAAAATT 540
QY	181 GlyGlnArgLeuGlnLeuLeuAsnTYrSerAsnSerThrGlnIleArgValAlaArgIle 200
DB	541 GGACAAACGCTTGAACCTAATAATTAATTTCCAAATCTACAGAAATACGCGTAGGCCAATT 600
QY	201 GlnGluIleCysGlyArgArgMetAsnValSerIleThrLysLysAspPheProGluSer 220
DB	601 CAAAGAAATATGAGCAAGCAAGATGAATGATATCATACAAAGAAACACTTCCCGAATGC 660
QY	221 LeuProAspAlaAspAspArgGlnValPheSerSerGlySerGlnTYrTrpIleAsp 240
DB	661 CTTCGAGATGCGAGATGACACACAGACAAGCTTACCTCGAGATTCACAAATATGGATAGAC 720
QY	241 GluGlySerPhePheIlePheProValGlyPheAlaAlaValAsnGlyTYrGlnLeuAsn 260
DB	721 GAGGGAAGCTTCTCATATTTCTCTGTGGATTTCGACAGCTCAATGATATCAACTAAAT 780
QY	261 AlaLysLysGluTYrIleGluHisThrAsnLysIleAlaGlnAlaIleLysAsnGlyGlu 280
DB	781 GCGAAAAAGGATATATTCAGCACACAAATTAATTTGCTCAACCATTAATAAATGGAGAA 840
QY	281 AsnProArgTYrAspSerAspAspValThrPheAspGlnLeuAlaLysAspProIleAsp 300
DB	841 AATCCAAAGTATGACTCAGACAGCGACATTTGATTCATTTAGCAAAAGATCCAAATTTGAT 900

QY	301	Prometllerprbrqglvsvallysvaiglylnphspegileuileaspproleuala	320
Db	901	CCCATGATTTGGAGAAAGTTAAAGTTGGACAAAAGTTTAAAGCTCAGTCAGACCCCTTGSGT	960
QY	321	GlncIlnpheasnslnuuhisvalalaserlleuulysphecyslystrhgulglytyr	340
Db	961	CAGCAATTCATTAACCTCCACGTCGCTTCGATTCTCAAAATTTTGCAAAACGTAAGAGATAT	1020
QY	341	LeuilevalglmetasprglyprobspalaaleuglnuspserPheProIlehisIleasn	360
Db	1021	CTTATTGTGGCAAGATGGTGTCCAGATGCACCTTGAAAGACAGTTTTCTATTCAATCAAT	1080
QY	361	AsntrhpheMetPheProValIGlytyrAlaGluIlystryAsnleuGluLeuValProPro	380
Db	1081	AATACATTTTATGTTCCACAGTTGGTTATGCGGAAAAGATATAATTGGAACTTGTCGCGCA	1140
QY	381	AspGluDbhelysgllythrPheArgtyrPaspGluItyrLeuGluIyGluSerAlaGluThr	400
Db	1141	GATGAGTTCAAGAGACATTCAGATGGGATGAAATACTTGGAGAAAGAAATCTGCAGAAACC	1200
QY	401	LeuProleuaspLeuPheIyysPrometProserIln-GlnAArgleuaspIyysPheIyysVa	420
Db	1201	CTACCGGTGACTGTTGTTCAAGCCCAATCCCTTCCGA-AGAGGATTTAGCAAAATTTAAAGT	1259
QY	420	IleleuIleeserIysArqValIGlyLeuArgleuGluAlaIlaIaspMetCysGluAsnGI	440
Db	1260	AATTCGTGATTTCCAAACGGGTAGAGACTACGCGCTTGAAAGCTGCTGACANTGTGAAAATCA	1319
QY	440	nPheIleCysProAlaThrValIyysSerValHisIGlyArgleuIleasnValAsnPheas	460
Db	1320	GTTTATTTTGTGCCACTACAGTGAATACATTCATCAGGAAGACGTGATTAATGTCAATTTGCA	1379
QY	460	pgIlyTTPaspGluDbpheasprGluLeuTyAspValIaspSerHisAspIleuProIle	480
Db	1380	CGGCTGGGATGCMAAATTTGATGAACGTATGATGATGAGATGCATCCCATGATATTTACCGAT	1439
QY	480	egIlyTTPCysGluAlaHisSerTyValleuGlnProProIyysIystryAsnTyR	498
Db	1440	AGATGGTGGTGATACCGGCACATTTATGTTCTCAACACCTCCGAAAAGATCAACATAT	1494
RESULT 3			
ABL52516			
ID	ABL52516	standard; DNA; 1497 BP.	
XX	AC	ABL52516;	
XX	XX		
DT	17-JUL-2002	(first entry)	
XX	XX		
De	C. elegans mutant lin-61	DNA sequence lin-61(n3447) SEQ ID NO:74.	
XX	XX		
KW	Caenorhabditis elegans; lin-8; lin-56; lin-61;	tumour suppressor;	
XX	XX	cell proliferation; nematode; cancer; mutant; gene; ds.	
XX	DS	Caenorhabditis elegans.	
XX	Key	location/Qualifiers	
FT	CDS	1..1497	
FT		/+tag= a	
FT		/product= "LIN-61(n3447) protein"	
XX	XX		
PN	WO200194545-A2.		
XX	PD		
XX	13-DEC-2001.		
XX	01-JUN-2001; 2001WO-US17909.		
XX	02-JUN-2000; 2000US-208802P.		
XX	PR		
XX	PA	(MASI) MASSACHUSETTS INST TECHNOLOGY.	
XX	PI	Horvitz HR, Davison EM, Lu X;	
XX	XX		

DR	rpi-1; 2002-401590/43.
DR	P-PSDB; ABB786598.
XX	
PT	New Caenorhabditis elegans nucleic acid involved in tumor suppressor pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
PT	diagnosing and treating cell proliferative diseases such as cancer
XX	
XX	Claim 27; Page 113; 116pp; English.
CC	The present invention describes a substantially pure nematode C. elegans (Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide (see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide (see ABB78650), where the polypeptides comprise at least 130, 110, 130 CC contiguously amino acids of the 386, 322 and 498 amino acid sequences given CC in ABB78648 to ABB78650, and modulate cell proliferation. (I) is useful CC for cytoskeletal activity, and can be used in gene therapy. (II) is useful CC for modulating proliferation of a cell, and for identifying the compound that CC modulates cell proliferation. (II) can be used for diagnosing an animal CC (preferably, human) for the presence of the cell proliferation disease, CC or an increased chance of developing the disease, by measuring lin-8, CC lin-56 or lin-61 nucleic acid expression in a sample obtained from the CC animal, by measuring LIN-8, LIN-56 or LIN-61 polypeptide in the sample CC using Ab or measuring amount of lin-8, lin-56, lin-61 mRNA in the sample CC (I) is also useful for identifying the compound that modulates cell CC proliferation. (I) and its mutants are useful for understanding cell CC proliferative diseases including cancer, as well as in diagnosing and CC treating cell proliferative diseases. (I) is also useful for identifying CC tumour suppressors in other species such as mammals and may be used to CC identify therapeutic compounds. The present sequence encodes a CC C. elegans mutant LIN-61 protein from the present invention.
XX	
SQ	Sequence 1497 BP; 498 A; 272 C; 311 G; 416 T; 0 other:
Alignment Scores:	
Pred. No.:	0 Length: 1497
Score:	411.00 Matches: 497
Percent Similarity:	99.60% Conservative: 0
Best Local Similarity:	99.60% Mismatches: 1
Query Match:	82.53% Indels: 2
DB:	24 Gaps: 0
US-09-872-523-5 (1-498) x ABL52516 (1-1497)	
Oy	1 MetSerGluPheUeuLysLIleValAAGAlaAsnLYLSerASPArgLYSeuSPlys 20
Dd	1 ATGTCTGAATTTCTGAAANATGTCAACACTAACAAAAAATCGCAGACAATCATAG 60
Oy	21 ThrTYrLeuTPGIuSerTYrLeuHISgInPheGLuLYSGlyLysThSerPheIlePro 40
Dd	61 ACCTACTTGTGGGAATTCCTATTACATGCAGTTCGACAAAAGAAAACCTCTTCAATCCA 120
Oy	41 ValGluIlaPheAsnTrgAsnLeuTrrrValAsnPheASngILcYSValLYSGluLIyAL 60
Dd	121 GTTGAAACATTAATCAATGTAACCTTACAAGTTAATTTTAACGAATGCCGTAGAAGAGACTT 180
Oy	61 IlePheGluThrValAlaHisASPTrYrAspLYSAsnCYSAspSerIleGlnValArgrTrp 80
Dd	181 ATCTTCGAAAACAGGGGTCATGATTTGACAAACAACTCGCATTCGATTCAAATCAGATCG 240
Oy	81 PheAlarGlIleGluLYSValCYSGlyTYrArgVALleuAlaGlnPheIleGlyIlaAsp 100
Dd	241 TTGTGCAGAAATGAAAAAGTTTGGGATCACAGGTTCGTGCAGTTTATCGGACCTGCAC 300
Oy	101 ThrLYSPHeTrPLEuAsnIleLeuSerASPmetPheGlyLEuAlaAsnAlaIamet 120
Dd	301 ACCGAATTTTGGCTCATATATTTTATTCGGACAGARATGTTTGTTGGCAAACGCCCATNG 360
Oy	121 SeraspTroksAnMetasplySLileValTYrAlaProProlEuAlalIeasnGLuclutyr 140
Dd	361 AGTAGTCCCAATFATGTATAAATTTGATATATGCTCCGCCCTTGCAATCAACGAAAGATAC 420
Oy	141 GlInAsnspHetValasnrTYrValaAsnaSCysIIeaSpGlyGluIleValGelYgInTrp 160

Db 421 CAAATGATATGTAATTATGTAATTAATTCATTGATGCGCAAAATCGCGCAAACT 480
 QY SerLeuSerProLysPheAspGluLysAlaLeuSerLysHisArgPheLysVal 180
 Db 481 TCGCTGTCTCCAAATTCATGAGGAGGAGGCTCTTAAGCAAGATGATTCAAAGTT 540
 QY 181 G1yGlnArgLeuGluLeuLeuAsnTyrSerAsnSerThrGluIleArgValAlaArgIle 200
 Db 541 GGCACAACTCTGAATATTAATTTCCAAATTCACGCAATACGCGTACGCCGAATT 600
 QY 201 G1nGluIleCysGlyArgArgMetAsnValSerIleThrLysLysAspPheProGluSer 220
 Db 601 CAAGAAATATGTGAGACGAGCAATGAATGTATCATCCAAAGCAAACTTCCCGAATCG 660
 QY 221 LeuProAspAlaAspAspAspArgGlnValPheSerSerGlySerGlnTyrTrpIleAsp 240
 Db 661 CTTCAGATTCAGATGACACACACACAGCTTTAGCTGAGATCTCAATATTTGGATAGAC 720
 QY 241 G1nGlySerPhePheIlePheProValGlyPheAlaAlaValAsnGlyTyrGlnLeuAsn 260
 Db 721 GAGGGAAGCTTCTCATATTTCTCTGTTGGATTTGCAGCATCATGATATCAACTAAT 780
 QY 261 AlaLysLysGluTyrIleGluHisThrAsnLysIleAlaGlnAlaIleLysAsnGlyGlu 280
 Db 781 GCGAAAAAGCAATATATATGAGCACACAAATAAATTTGCTCAACCAATAAAAATGGAGAA 840
 QY 281 AsnProArgTyrAspSerAspAspValThrPheAspGluLeuAlaLysAspProIleAsp 300
 Db 841 AATCCAAAGATATGACACACACACGCTCAATTTGATCAATTAGCAAAAGATCCCAATTGAT 900
 QY 301 PrometIleTyrArgLysValLysValGlyGlnLysPheGluLeuIleAspProLeuAla 320
 Db 901 CCCATGATTTGGAGAAAAGTTAAGGTTGGCAAAAAGTTGAGCTCATCCACCCTTGCGCT 960
 QY 321 G1nGlnPheAsnAsnLeuHisValAlaSerIleLeuLysPheCysLysThrGluGlyTyr 340
 Db 961 CAGCAATTCATATTAACCTCCACGCTGCTGATTTCTCAAAATTTGGCAAACTGAGGATAT 1020
 QY 341 LeuIleValGlyMetAspGlyProAspAlaLeuGluAspSer-PheProIleHisIleAs 360
 Db 1021 CTATATGTGGAAATGATGTCACAGATGCACATTGAAGACAA-TTTTCTTATTCATATCA 1079
 QY 360 nasnthrPheMetPheProValGlyTyrAlaGluLysTyrAsnLeuGluLeuValProPr 380
 Db 1080 TATATACATTTATGTTCCAGTTGCTATATCGGAAAAGTAAATTTGCAACTGTTCCGCC 1139
 QY 380 AspGluPheLysGlyThrPheArgTrpAspGluTyrLeuGluLysGluSerAlaLuh 400
 Db 1140 AGATGAGTTCAAAGCAATTCAGATGGATGAATCTTGAGAAAAGAAATCTCCAGAAAC 1199
 QY 400 rLeuProLeuAspLeuPheLysProMetProSerGlnGluArgLeuAspLysPheLysVal 420
 Db 1200 CCGACCGCTTGACTTTCACGCCAATGCTTCCCAAGACATTTGACAAATTTTAAGT 1259
 QY 420 11leuIleSerLysArgValGlyLeuArgLeuGluAlaAlaAspMetCysGluAsnG1 440
 Db 1260 AATTCGATTTCCAAACGGGTAGACTACGCTTGAAAGCTGCGACATGTGTGAATAATA 1319
 QY 440 nPheIleCysProAlaThrValLysSerValHisGlyArgLeuIleAsnValAsnPheAs 460
 Db 1320 GTTATATTTCCAGCTACAGTGAATCACTTCATGGAACATCGATTAATTCATTTGCA 1379
 QY 460 pGlyTyrAspGluGluPheAspGluLeuTyrAspValAspSerHisAspIleLeuProI1 480
 Db 1380 CCGCTGGATGCAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1439
 QY 480 eGlyTyrCysGluAlaHisSerTyrValLeuGlnProPolysLysTyrAsnTyr 498
 Db 1440 AGATGTGTGAAGCGCACAGTTATGTCTACAACCTCGAAAAAGTACAACATAT 1494

RESULT 4

ABL52517
 ID ABL52517 standard; DNA; 1497 BP.

XX XX
 AC ABL52517;
 XX
 DT 17-JUL-2002 (first entry)
 XX
 DE C. elegans mutant lin-61 DNA sequence lin-61(n3624) SEQ ID NO:75.
 XX
 KW Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;
 XX cell proliferation; nematode; cancer; mutant; gene; ds.
 OS
 FH Caenorhabditis elegans.
 FT
 FT CDS Location/Qualifiers
 FT 1..1497
 FT /tag= a
 FT /product= "LIN-61(n3624) protein"
 PN WO200194545-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 01-JUN-2001: 2001MO-US17909.
 XX
 PR 02-JUN-2000: 2000US-208802P.
 XX
 PA (MAS1) MASSACHUSETTS INST TECHNOLOGY.
 PI Horvitz HR, Davison EM, Lu X;
 XX
 DR WPI; 2002-401590/43.
 XX
 DR P-PSDB; ABB78699.
 PT
 PT New Caenorhabditis elegans nucleic acid involved in tumor suppressor
 PT pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
 PT diagnosing and treating cell proliferative diseases such as cancer
 PS
 PS Claim 27; Page 113-114; 116pp; English.
 XX
 CC The present invention describes a substantially pure nematode C. elegans
 CC (Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide
 CC (see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide
 CC (see ABB78650), where the polypeptides comprise at least 130, 110, 130
 CC contiguous amino acids of the 386, 322 and 498 amino acid sequences given
 CC in ABB78648 to ABB78650, and modulate cell proliferation. (I) has
 CC cytostatic activity, and can be used in gene therapy. (I) is useful for
 CC modulating proliferation of a cell, and for identifying the compound that
 CC modulates cell proliferation. (I) can be used for diagnosing an animal
 CC (preferably, human) for the presence of the cell proliferation disease,
 CC or an increased chance of developing the disease, by measuring lin-8,
 CC lin-56 or lin-61 nucleic acid expression in a sample obtained from the
 CC animal, by measuring LIN-8, LIN-56 or LIN-61 polypeptide in the sample
 CC using Ab or measuring amount of lin-8, lin-56, lin-61 mRNA in the sample.
 CC (I) is also useful for identifying the compound that modulates cell
 CC proliferation. (I) and its mutants are useful for understanding cell
 CC proliferative diseases including cancer, as well as in diagnosing and
 CC treating cell proliferative diseases. (I) is also useful for identifying
 CC tumor suppressors in other species such as mammals and may be used to
 CC identify therapeutic compounds. The present sequence encodes a
 CC C. elegans mutant LIN-61 protein from the present invention.
 XX
 XX Sequence 1497 BP; 497 A; 271 C; 312 G; 417 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 0 length: 1497
 Score: 397.00 Matches: 497
 Percent Similarity: 99.608
 Best Local Similarity: 99.608
 Query Match: 79.728 Mismatches: 0
 DB: 24 Indels: 2
 Gaps: 0

US-09-872-523-5 (1-498) x ABL52517 (1-1497)

QY 1 MetSerGluPheLeuLysIleValArgAlaAsnLysLysSerAspArgLysLeuAspLys 20

```

|||||
Db 1 ATGCTCAATTTTCGAAAAATTCAGAGCTAACAAAAATCGACACAAAACTCGATAG 60
Oy 21 ThrTyrLeuTrpGluSerTyrLeuHisGlnPheGluTyrGlyLysThrSerPheIlePro 40
Db 61 ACCACTTGTGGGAATCCTATTTCATCATCAGTTCGAGAAAGAAAACTCTTCATTCCCA 120
Oy 41 ValGluAlaPheAsnArgAsnLeuThrValAsnDheAsnGluCysValLysGluGlyVal 60
Db 121 GTTGAACCATTCATCATCTTAACCTTACAGTAAATTTTAAAGCAATCCGTCGAAGAGAGATT 180
Oy 61 IlePheGluTrpThrValHisAspTyrAspLysAsnGlyAspSerIleGlnValArgTrp 80
Db 181 ATCTTCGAACAGTGGCCATGATTATGACAAAGACTCGATTCGATTCAAGTCAGATGG 240
Oy 81 PheAlaArgIleGluLysValCysGlyTyrArgValIleuAlaGlnPheIleGlyAlaAsp 100
Db 241 TTTCGACGAATTTGAAAAAGTTTGGGAGTACAGACTTCGGCTCGAGTTATTCGGAGCTGAC 300
Oy 101 ThrLysPheTrpLeuAsnIleLeuSerAspAspMetPheGlyLeuAlaAsnAlaMet 120
Db 301 ACGAATTTTGGCTCAATATTTTATCGGACGATATGTTGGTTGGCAAAACGCCGCAATG 360
Oy 121 SerAspProAsnMetAspLysIleValTyrAlaPro-proLeuAlaIleAsnGluGluTyr 140
Db 361 AGTATCCCAATATGAGTAAATGTAATATGCTTC -GCCGCTTCGAATCAACGAAGATA 419
Oy 140 rGlnAsnAspMetValAsnTyrValAsnAsnCysIleAspGlyGluIleValGlyGlnTh 160
Db 420 CCAAAATATATGTGAAATTTATGTAATAATTTGCATTGATGGCAAAATCGTCGGCAAC 479
Oy 160 rSerLeuSerProLysPheAspGluGlyLysAlaLeuLeuSerLysHisArgPheLysVa 180
Db 480 TTCCGCTCTCCAAATTCGATGAAGGAGCGCTCCCTAAGCAACGACTTCCTTCAAGT 539
Oy 180 IGLYlnArgLeuGluLeuLeuAsnTyrSerAsnSerThrGluIleArgValAlaArgI 200
Db 540 TGGACACGCTTGTGACATTAATTAATTCATTCATCTGAATATCCGTCAGCCGAT 599
Oy 200 eGlnGluIleCysGlyArgArgMetAsnValSerIleThrLysLysAspPheProGlu 220
Db 600 TCAAGAAATATGTGGACGACGACGAATGATGTATCTATCAAAAGAAAGACTTCCCGAATC 659
Oy 220 rLeuProAspAlaAspAspArgGlnValPheSerSerGlySerGlnTyrTrpIleAs 240
Db 660 GCTTCACATGTCAGATGACAGACAGCAAGCTTTAGCTCTGGATCTCATATTTGGATAGA 719
Oy 240 pGluGlySerPhePheIlePheProValGlyPheAlaValaValaAsnGlyTyrGlnLeuAs 260
Db 720 CGAGGGAAGCTTCTTCATATTTCCGTTGGATTTCACACAGTCATGGAATCACTATAA 779
Oy 260 nAlaLysLysGluTyrIleGluHisThrAsnLysIleAlaGlnAlaIleLysAsnGlyG 280
Db 780 TGGCAAAAAGGAATATATTTAGACACACAATAATTAATTCCTCAAGCAATAAAAAAGGAGA 839
Oy 280 uAsnProArgTyrAspSerAspAspValThrPheAspGlnLeuAlaLysAspProIleAs 300
Db 840 AAATCCAAAGATATGACTCAGACGACGATTCATCAATTTAGCAAAAAGATCCAATTTGA 899
Oy 300 ProMetIleTrpAlaGlyValLysValGlyGlnLysPheGluLeuIleAspProLeuAl 320
Db 900 TCCCATGATTTGGACAAAAGTTAAGCTTGGACAAAAGTTTGAGCTCATCGACCCCTTGGC 959
Oy 320 aGlnGlnPheAsnAsnLeuHisValAlaSerIleLeuLysPheCysLysThrGluGlyTyr 340
Db 960 TCAGCAATTCATTAACCTCCACGCTCGCTTCGATTCCTCAAAATTTTCAAAAAGAGAGATA 1019
Oy 340 rLeuIleValGlyMetAspGlyProAspAlaLeuLysAspSerPheProIleHisIleAs 360
Db 1020 TCTTATTTGGGAAATGGATGGTCAGATGCACCTTGAAGACACTTTTCTATTCATATCA 1079
Oy 360 nAsnThrPheMetPheProValGlyTyrAlaGluTyrTrpAsnLeuGluLeuValProPr 380

```

```

Db 1080 TAATACATTTATGTTCCCATGTTGGTATGGGAAAGATATAATTTGGAACTGTTCGCC 1139
Oy 380 oAspGluPheLysGlyThrPheArgTrpAspGluTyrLeuGluLysGluSerAlaGluTh 400
Db 1140 AGATGAGCTTCAAGGAACATTCAGATGGATGATATCTTGAGAGAAAGAAATCTCGAAGAC 1199
Oy 400 rLeuProLeuAspLeuPheLysProMetProSerGlnGluArgLeuAspLysPheLysVa 420
Db 1200 CCTACCGCTTGACTGTTTCAGGCCAATGCCCTTCCCAAGAGAGATTACACAAATTTAAGCT 1259
Oy 420 lIleLeuIleSerLysArgValGlyLeuArgLeuGluAlaAlaAspMetCysGluAsnG 440
Db 1260 AATTCGATTTCCAAACGGGAGAGACTACGCTTGAAGCTGTCATCATGTGTGAAAAATCA 1319
Oy 440 nPheIleCysProAlaThrValLysSerValHisGlyArgLeuIleAsnValAsnPheAs 460
Db 1320 GTTATATTTGTCACACTCAGAGAAATCAGTTCAAGAGAACTGATTAATGTCATTTTCGA 1379
Oy 460 pGlyTrpAspGluGluPheAspGluLeuTyrAspValAspSerHisAspIleLeuProI 480
Db 1380 CGGCTGGAGTGAAGAATTTGATGAACTGTATGATGTGACTCCCATGATATTTTACCGAT 1439
Oy 480 eGlyTrpCysGluAlaHisSerTyrValLeuGlnProProLysLysTyrAsnTyr 498
Db 1440 AGATGCGTGTGAACGACGACAGTATGTTCTACAACTCCGAAAAAGTACAACTAT 1494

RESULT 5
ID ABL52518 standard; DNA; 2307 BP.
XX ABL52518;
AC ABL52518;
XX 17-JUL-2002 (first entry)
DE C. elegans lin-61 DNA sequence SEQ ID NO:76.
XX Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;
KM cell proliferation; nematode; cancer; mutant; gene; ds.
XX Caenorhabditis elegans.
OS WO200194545-A2.
PN 13-DEC-2001.
XX 01-JUN-2001; 2001WO-US17909.
PE 02-JUN-2000; 2000US-208802P.
PR (MASI ) MASSACHUSETTS INST TECHNOLOGY.
PA Horvitz HR, Davison EM, Lu X;
PI WPL; 2002-401590/43.
DR The present invention describes a substantially pure nematode C. elegans
XX pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
PT diagnosing and treating cell proliferative diseases such as cancer
XX
XX Disclosure; Page 114-115; 116pp; English.
XX The present invention describes a substantially pure nematode C. elegans
XX (Caenorhabditis elegans) nucleic acid (i) encoding a LIN-8 polypeptide
XX (see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide
XX (see ABB78650), where the polypeptides comprise at least 130, 110, 130
XX contiguous amino acids of the 386, 322 and 498 amino acid sequences given
XX in ABB78648 to ABB78650, and modulate cell proliferation. (i) has
XX cytosolic activity, and can be used in gene therapy. (i) is useful for
XX modulating proliferation of a cell, and for identifying the compound that
XX modulates cell proliferation. (i) can be used for diagnosing an animal
XX (preferably, human) for the presence of the cell proliferation disease,
XX or an increased chance of developing the disease, by measuring lin-8,
XX lin-56 or lin-61 nucleic acid expression in a sample obtained from the

```

CC animal, by measuring LIN-8, LIN-56 or LIN-61 polypeptide in the sample.
 CC using Ab or measuring amount of lin-8, lin-56, lin-61 mRNA in the sample.
 CC (I) is also useful for identifying the compound that modulates cell
 CC proliferation. (I) and its mutants are useful for understanding cell
 CC proliferative diseases including cancer, as well as in diagnosing and
 CC treating cell proliferative diseases. (I) is also useful for identifying
 CC tumour suppressors in other species such as mammals and may be used to
 CC identify therapeutic compounds. The present sequence represents a
 CC C. elegans lin-61 DNA sequence from the present invention.

Sequence 2307 BP; 740 A; 393 C; 466 G; 708 T; 0 other:

Alignment Scores:

Pred. No.:	5.21e-285	Length:	2307
Score:	278.00	Matches:	278
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	55.82%	Indels:	0
DB:	24	Gaps:	0

US-09-872-523-5 (1-498) x ABL52518 (1-2307)

QY 150 AaNCyIIlaSpGIyLuIleValIGlyInThSerIeuSerProLyshPhaSpGIuGly 169
 Db 1138 AATTCATTGATGGCAAAATCGCGCAAACTCGCTGCTCCAAAATTCGATGAAGG 1197
 QY 170 LysAlaLeuIeuSerYshIsArGPhelYsValIGlyInArGleuGIuIeuAaNTyr 189
 Db 1198 AAGGCTCTCTTAAGCAGCATCGTTCCAAAGTTGGCAACGCTTTAACTATTAAATTAT 1257
 QY 190 SerAsnSerThrGIuIleArGyValAlaArGIleGIuIleCysGIyArGArGMeIasn 209
 Db 1258 TCCAATTTACTGAATACGCGGATGCGCAATTCAGAAATATGTGACAGACAAATGAAT 1317
 QY 210 ValSerIleThrLyshYsAspPherProGIuSerIeuProAspAlaAspAspArGIn 229
 Db 1318 GATTCATATCAAAAGAAACCTTCCGAAATCGCTTCAGATGCAATGACGACAGACAA 1377
 QY 230 ValPheSerSerGIySerGIuNTyrTrPIleAspGIuGIySerPheheIlePheProVal 249
 Db 1378 GTCTTAGCTCTGATCTCAATATTGGATAGACGAGGAAGCTTCTTCATATTCTCTGTT 1437
 QY 250 GlyPheAlaAlaValaAsnGIyTyrGIuIeuAsnAlaLysGIuTyrIleGIuIshThr 269
 Db 1438 GGATTGCGACAGTCATGATGATCAACTAAATGCGAAAGAAATATATATTGAGCACACA 1497
 QY 270 AsnLyIleAlaGIuAlaIleLysAsnGIyGIuAsnProArGTrAspSerAspArVal 289
 Db 1498 AATAAATTCCTCAACCAATATAAAATGAGAAATCCAAAGATATGACTCAGACGACGTC 1557
 QY 290 ThrPheAspGIuIeuAlaLysAspProIleAspProMetIleTrpArGIyValLysVal 309
 Db 1558 ACATTTGATCATTTAGCAAAAGATCCCAATTTGATCCCATGTTTGGAGAAAGTTAAGGTT 1617
 QY 310 GIyGIuLysPheGIuIeuIleAspProLeuAlaGIuInGIuPheAsnAsnLeuHIsVala 329
 Db 1618 GGCAGAAAGTTTGAGCTCATCGACCCCTTGGCTCAGCAATCAATACCTCCACGTCGCT 1677
 QY 330 SetIleLeuLysPheCysLyshThrGIuGIyTyrIleuIleValaGIyMetAspGIyProAsp 349
 Db 1678 TCGATTCTCAAAATTTTGGCAAACTGGAAGATATCTTATGTGTGGAAATGATGGTCCAGAT 1737
 QY 350 AlaLeuGIuAspSerPheProIleHIsIleAsnAsnThrPheMetPheProValGIyTyr 369
 Db 1738 GCACCTGAAGACAGTTTCTCTATTATATCAATTAATATATATTTCCAGTTGGTTAT 1797
 QY 370 AlaGIuLysTyrAsnLeuGIuLeuValProProAspGIuPheLysGIyThrPheArGTrp 389
 Db 1798 GCGGAAAGATATATTTGGAACCTGTTCGCCGAGATGAGTTCAAAGAACATTCAGATGG 1857
 QY 390 AspGIuTyrLeuGIuLysGIuSerAlaGIuThrLeuProLeuAspLeuPheLysProMet 409
 Db 1858 GATGAATACTTGGAGAAAGATCTGCAGAAACCTTCGCTTGACTTGTTCAGCCCAATG 1917

QY 410 ProSerGIuIuArGleuAspLysPheLysValIleLeuIleSerYsArGVal 427
 Db 1918 CCTTCCCAAGAGATTAACAAATTTAAGTATTTCTGATTTCCAAACGGTT 1971

RESULT 6

ABL52519

ID ABL52519 standard; DNA; 2307 BP.

XX ABL52519;

AC 17-JUL-2002 (first entry)

DE C. elegans lin-61(sy223) DNA sequence SEQ ID NO:77.

XX Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;

KW cell proliferation; nematode; cancer; mutant; gene; ds.

XX Caenorhabditis elegans.

XX MO200194545-A2.

XX 13-DEC-2001.

PD 01-JUN-2001; 2001WO-US17909.

XX 02-JUN-2000; 2000US-208802P.

XX (MAST) MASSACHUSETTS INST TECHNOLOGY.

XX Horvitz HR, Davison EM, Lu X;

XX WPI; 2002-401590/43.

PT New Caenorhabditis elegans nucleic acid involved in tumor suppressor

PT pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for

PT diagnosing and treating cell proliferative diseases such as cancer

XX Disclosure; Page 115; 116pp; English.

CC The present invention describes a substantially pure nematode C. elegans
 CC (Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide
 CC (see ABB78648), LIN-56 polypeptide (see ABB76649), or LIN-61 polypeptide
 CC (see ABB78650), where the polypeptides comprise at least 130, 110, 130
 CC contiguous amino acids of the 386, 322 and 498 amino acid sequences given

CC in ABB76648 to ABB78650, and modulate cell proliferation. (II) has
 CC cytosolic activity, and can be used in gene therapy. (I) is useful for
 CC modulating proliferation of a cell, and for identifying the compound that

CC modulates cell proliferation. (I) can be used for diagnosing an animal
 CC (preferably, human) for the presence of the cell proliferation disease,
 CC or an increased chance of developing the disease, by measuring lin-8,
 CC lin-56 or lin-61 nucleic acid expression in a sample obtained from the

CC animal, by measuring amount of lin-8, lin-56, lin-61 mRNA in the sample.
 CC (I) is also useful for identifying the compound that modulates cell
 CC proliferation. (I) and its mutants are useful for understanding cell
 CC proliferative diseases including cancer, as well as in diagnosing and

CC treating cell proliferative diseases. (I) is also useful for identifying
 CC tumour suppressors in other species such as mammals and may be used to
 CC identify therapeutic compounds. The present sequence represents a
 CC C. elegans lin-61(sy223) DNA sequence from the present invention.

XX

XX

XX

Sequence 2307 BP; 741 A; 393 C; 465 G; 708 T; 0 other:

Alignment Scores:

Pred. No.:	5.21e-285	Length:	2307
Score:	278.00	Matches:	278
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	55.82%	Indels:	0
DB:	24	Gaps:	0

US-09-872-523-5 (1-498) x ABL52519 (1-2307)

QY 150 AsnCysIleaspglygluilevalglycInThrSerLeuSerProLysPheaspgluGly 169
 |||||
 Db 1138 AATTCATGATGCGCAAAATCGTCGGCCAAACTCGCTGCTCCAAAATTCGATGAAGGG 1197
 QY 170 LysAlaLeuLeuSerLysHisArgPheLysValGlyGlnArgLeuGluLeuLeuAsnTyr 189
 |||||
 Db 1198 AAGGCTCTCTTAACCAAGCATCGTTTCAAGTTGGACAGCTCTTGAATTAATTAAT 1237
 QY 190 SerAsnSerThrGluIleArgValAlaArgIleGlnGluIleCysGlyArgArgMetAsn 209
 |||||
 Db 1258 TCCAAATTCATGAAATACGCGTACCGCGAATTCAGAAATATGTGGACGACGATGAT 1317
 QY 210 ValSerIleThrLysLysAspPheProGluSerLeuProAspAlaAspAspArgGln 229
 |||||
 Db 1318 GTATCTATCACAAAGAAAGAACTTCCGAAATCGCTTCAGATGACAGACAGACAA 1377
 QY 230 ValPheSerSerGlySerGlnTyrTrpIleAspGluGlySerPhePheIlePheProVal 249
 |||||
 Db 1378 GTCTTTAGCTCTGGAATCTCAATATTGGATAGACGAGGAGACCTTCTTCATATTCTGCT 1437
 QY 250 GlyPheAlaAlaValAsnGlyTyrGlnLeuAsnAlaLysLysGlyTyrIleGluHisThr 269
 |||||
 Db 1438 GGATTTGCGACGACGATCATGATTCGATCACTAAATGCGAAAGGAAATATATTGACGACACA 1497
 QY 270 AsnLysIleAlaGlnAlaIleLysAsnGlyLysAsnProArgTyrAspSerAspAspVal 289
 |||||
 Db 1498 AATAAAATGCTCAGCAAAATGAGAAAAATGCAAAATCCAAATATGACTCAGACGACGCTC 1557
 QY 290 ThrPheAspGlnLeuAlaLysAspProIleAspProMetIleTyrPargLysValLysVal 309
 |||||
 Db 1558 ACATTTGATTCATTAATGCAAAAGATCCCAATTCATCCCATGATTGGAGAAAAGTTAAGTT 1617
 QY 310 GlyGlnLysPheGluLeuIleAspProLeuAlaGlnGlnPheAsnAsnLeuHisValAla 329
 |||||
 Db 1618 GGACAAAAGTTGAGCTCATGCGACCCCTGGCTCAGCAATTAACCTCCACGCTCGCT 1677
 QY 330 SerIleLeuLysPheCysLysThrGlnGlyTyrLeuIleValGlyMetAspGlyProAsp 349
 |||||
 Db 1678 TCGATTCCTCAAAATTTGCAAAATGAAAGATATCTTATTGTGGAGATGATGCTCCAGAT 1737
 QY 350 AlaLeuGluAspSerPheProIleHisIleAsnAsnThrPheMetPheProValAlaGlyTyr 369
 |||||
 Db 1738 GCACCTGGAAGACAGTTTCCATTCATTCATCAATATACATTTATGTTCCAGTTGTTAT 1797
 QY 370 AlaGluLysTyrAsnLeuGluLeuValProProAspGluPheLysGlyThrPheArgTrp 389
 |||||
 Db 1798 GCGGAAAAGTAAATTTGGAACCTTGTCGCCAGATGAGTTCAAAAGGAAACATTCAGATGG 1857
 QY 390 AspGluTyrLeuGluLysGluSerAlaGlnThrLeuProLeuAspLeuPheLysProMet 409
 |||||
 Db 1858 GATGAAATCTTGAGAAAAGAACTGCGAAGAACCTTACCGGCTTGACTTGTTCAGGCCAATG 1917
 QY 410 ProSerGlnGluArgLeuAspLysPheLysValIleLeuIleSerLysArgVal 427
 |||||
 Db 1918 CCTTCCCAAGAGAGATTGACAAATTTAAGTAATTCGATTTCGAAACGGGTT 1971
 RESULT 7
 ABL52520
 ID ABL52520 standard; DNA; 2307 BP.
 NC ABL52520;
 XX
 DT 17-JUN-2002 (first entry)
 DE C. elegans lin-61(n3635) DNA sequence SEQ ID NO: 78.
 XX
 XX Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;
 KM cell proliferation; nematode; cancer; mutant; gene; ds.
 OS Caenorhabditis elegans.
 XX
 PN MO200194545-A2.

XX 13-DEC-2001.
 PD
 XX 01-JUN-2001; 2001MO-US17909.
 PF
 XX 02-JUN-2000; 2000US-208802P.
 PR
 XX (MAST) MASSACHUSETTS INST TECHNOLOGY.
 PA
 XX Horvitz HR, Davison EM, Lu X;
 PI WPI; 2002-401590/43.
 DR
 XX
 XX
 PT New Caenorhabditis elegans nucleic acid involved in tumor suppressor
 pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
 PT diagnosing and treating cell proliferative diseases such as cancer
 XX
 XX Disclosure; Page 116; 116pp; English.
 PS
 CC The present invention describes a substantially pure nematode C. elegans
 CC (Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide
 CC (see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide
 CC (see ABB78650), where the polypeptides comprise at least 130, 110, 130
 CC contiguous amino acids of the 386, 322 and 498 amino acid sequences given
 CC in ABB78648 to ABB78650, and modulate cell proliferation. (I) has
 CC cytosolic activity, and can be used in gene therapy. (I) is useful for
 CC modulating proliferation of a cell, and for identifying the compound that
 CC modulates cell proliferation. (I) can be used for diagnosing an animal
 CC (preferably, human) for the presence of the cell proliferation disease,
 CC or an increased chance of developing the disease, by measuring lin-8,
 CC lin-56 or lin-61 nucleic acid expression in a sample obtained from the
 CC animal, by measuring LIN-8, LIN-56 or LIN-61 polypeptide in the sample.
 CC using Ab or measuring amount of lin-8, lin-56, lin-61 mRNA in the sample.
 CC (I) is also useful for identifying the compound that modulates cell
 CC proliferation. (I) and its mutants are useful for understanding cell
 CC proliferative diseases including cancer, as well as in diagnosing and
 CC treating cell proliferative diseases. (I) is also useful for identifying
 CC tumour suppressors in other species such as mammals and may be used to
 CC identify therapeutic compounds. The present sequence represents a
 CC C. elegans lin-61(n3635) DNA sequence from the present invention.
 XX
 SO Sequence 2307 BP; 741 A; 393 C; 465 G; 708 T; 0 other;
 Alignment Scores:
 Pred. No.: 5,21e-285 Length: 2307
 Score: 278.00 Matches: 278
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 55.82% Indels: 0
 DB: 24 Gaps: 0
 US-09-872-523-5 (1-498) x ABL52520 (1-2307)
 QY 150 AsnCysIleaspglygluilevalglycInThrSerLeuSerProLysPheaspgluGly 169
 |||||
 Db 1138 AATTCATGATGCGCAAAATCGTCGGCCAAACTCGCTGCTCCAAAATTCGATGAAGGG 1197
 QY 170 LysAlaLeuLeuSerLysHisArgPheLysValGlyGlnArgLeuGluLeuLeuAsnTyr 189
 |||||
 Db 1198 AAGGCTCTCTTAACCAAGCATCGTTTCAAGTTGGACAGCTCTTGAATTAATTAAT 1237
 QY 190 SerAsnSerThrGluIleArgValAlaArgIleGlnGluIleCysGlyArgArgMetAsn 209
 |||||
 Db 1258 TCCAAATTCATGAAATACGCGTACCGCGAATTCAGAAATATGTGGACGACGATGAT 1317
 QY 210 ValSerIleThrLysLysAspPheProGluSerLeuProAspAlaAspAspArgGln 229
 |||||
 Db 1318 GTATCTATCACAAAGAAAGAACTTCCGAAATCGCTTCAGATGACAGACAGACAA 1377
 QY 230 ValPheSerSerGlySerGlnTyrTrpIleAspGluGlySerPhePheIlePheProVal 249
 |||||
 Db 1378 GTCTTTAGCTCTGGAATCTCAATATTGGATAGACGAGGAGACCTTCTTCATATTCTGCT 1437

QY 250 GlyPheAlaAlaValaAsnGlyTyrGlnLeuAsnAlaLysGluTyrIleGlnHisThr 269
 |||||||
 Db 1438 GGATTGGACGACGTCAATGATATCAATAAATGGCAAAAGAAATATTTAGCAGCACAA 1497
 QY 270 AsnLysIleAlaGlnAlaIleLysAsnGlyLysAsnProAlaTyrAspSerAspVal 289
 |||||||
 Db 1498 AATAAATATGCTCAAGCAATAAATAATGAGAAATCCAAAGATATGACACACGCGC 1557
 QY 290 ThrPheAspGlnLeuAlaLysAspProIleAspProMetIleTyrArgLysVal 309
 |||||||
 Db 1558 AATTTGATCAATTCACCAAAAGATCCATATTTCCATATTTGGAGAAAGTTAGGTT 1617
 QY 310 GlnLysPheGlnLeuIleAspProLeuAlaGlnGlnPheAsnAsnLeuHisVala 329
 |||||||
 Db 1618 GGACAAAGTTGAGCTCATCGACCCCTTGGCTCAGCAATTCATATACCTCCACGTCGCT 1677
 QY 330 SerIleLeuLysPheCysLysThrGlnGlyTyrLeuIleValGlyMetAspGlyProAsp 349
 |||||||
 Db 1678 TCGATTCTCAAAATTTGGCAAACTGAAGATATCTTATGTGGATGATGTCGTCAGAT 1737
 QY 350 AlaLeuGluAspSerPheProIleHisIleAsnAsnThrPheMetPheProValGlyTyr 369
 |||||||
 Db 1738 GCACCTTGAGACAGTTTCTCATCATCAATTAATACATTAATGTTCACAGTTGGTTAT 1797
 QY 370 AlaGlnLysTyrAsnLeuGlnLeuValProAspGluPheLysGlyThrPheArgTyr 389
 |||||||
 Db 1798 GCGGAAAGTATATATTTGSACTGTGTCCGACAGATGATTCAAAGAACATTCAGATGG 1857
 QY 390 AspGlnTyrLeuGlnLysGlnSerAlaGlnThrLeuProLeuAspLeuPheLysProMet 409
 |||||||
 Db 1858 GATGATACTCTGAGAAAGAAATCTGCAGAAACCTACCGCTTGACTGTTCAGGCCAATG 1917
 QY 410 ProSerGlnGlnArgLeuAsnLysPheLysValIleLeuIleSerLysArgVal 427
 |||||||
 Db 1918 CCTTCCCAAGAGAGATTAGACAAATTTAAGTAAATTCGATTCCAAACGGGTT 1971

RESULT 8
 AAA31397/c
 ID AAA31397 standard; DNA; 344 BP.
 XX
 AC AAA31397;
 XX
 DT 05-JUL-2000 (first entry)
 XX
 DE Plant microsatellite marker #358.
 XX
 KW Plant microsatellite sequence; core repeat sequence: detection; probe;
 KW DNA polymorphism; genome mapping; physical mapping; fingerprinting;
 KW variety identification; genetic variability evaluation; primer; ss.
 XX
 OS Eucalyptus grandis.
 XX
 PN WO9967421-A1.
 XX
 PD 29-DEC-1999.
 XX
 PF 25-JUN-1999; 99WO-NZ00092.
 XX
 PR 25-JUN-1998; 98US-0105307.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD & FLETCHER.
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
 XX
 PI Havukkala IJ, Bloksberg LN, Glenn M;
 XX
 DR WPI; 2000-116958/10.
 XX
 PT New plant microsatellite markers and associated flanking species for
 PT the detection of polymorphic genetic markers -
 XX
 PS Claim 1; Page 182; 392pp; English.
 XX
 CC Sequences AAA31040-A32093 represent novel plant microsatellite sequences

CC and associated flanking species. The sequences comprise a central core
 CC repeat sequence, especially selected from the sequences AAA32094-A32096
 CC with left and right flanking sequences. The polynucleotide sequences
 CC can be used in the detection of DNA polymorphisms, in genome mapping,
 CC in physical mapping, in positional cloning of genes, in variety
 CC identification and in evaluation of genetic variability within and
 CC between plant tissues, populations, cultivars, species and species
 CC groups. They may also be used to design hybridization probes for
 CC oligonucleotide fingerprinting and library screening and to design
 CC primers for microsatellite-primed PCR. Microsatellite markers are
 CC useful to locate specific economically useful genes in plant genomes.
 XX
 SQ Sequence 344 BP; 72 A; 111 C; 93 G; 67 T; 1 other;
 Alignment Scores:
 Pred. No.: 6.05 Length: 344
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.81% Indels: 0
 DB: 21 Gaps: 0
 US-09-872-523-5 (1-498) x AAA31397 (1-344)
 QY 179 LysValGlnArgLeuGlnLeuLeu 187
 |||||||
 Db 273 AAGTAGACAGACGCTCGAACTTCTT 247

RESULT 9
 AAA31396/c
 ID AAA31396 standard; DNA; 350 BP.
 XX
 AC AAA31396;
 XX
 DT 05-JUL-2000 (first entry)
 XX
 DE Plant microsatellite marker #357.
 XX
 KW Plant microsatellite sequence; core repeat sequence: detection; probe;
 KW DNA polymorphism; genome mapping; physical mapping; fingerprinting;
 KW variety identification; genetic variability evaluation; primer; ss.
 XX
 OS Eucalyptus grandis.
 XX
 PN WO9967421-A1.
 XX
 PD 29-DEC-1999.
 XX
 PF 25-JUN-1999; 99WO-NZ00092.
 XX
 PR 25-JUN-1998; 98US-0105307.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD & FLETCHER.
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
 XX
 PI Havukkala IJ, Bloksberg LN, Glenn M;
 XX
 DR WPI; 2000-116958/10.
 XX
 PT New plant microsatellite markers and associated flanking species for
 PT the detection of polymorphic genetic markers -
 XX
 PS Claim 1; Page 182; 392pp; English.
 XX
 CC Sequences AAA31040-A32093 represent novel plant microsatellite sequences
 CC and associated flanking species. The sequences comprise a central core
 CC repeat sequence, especially selected from the sequences AAA32094-A32096
 CC with left and right flanking sequences. The polynucleotide sequences
 CC can be used in the detection of DNA polymorphisms, in genome mapping,
 CC in physical mapping, in positional cloning of genes, in variety
 CC identification and in evaluation of genetic variability within and
 CC between plant tissues, populations, cultivars, species and species
 CC groups. They may also be used to design hybridization probes for


```

OY      179 LysValGlyGlnArgLeuGluLeu 187
DB      276 AAGGTAGACAGCGTCTCGAACTTCTT 250

RESULT 12
AAA31392/c
ID      AAA31392 standard; DNA: 367 BP.
XX
AC      AAA31392;
XX
DT      05-JUL-2000 (first entry)
XX
DE      Plant microsatellite marker #353.
XX
KM      Plant microsatellite sequence; core repeat sequence; detection; probe;
KM      DNA polymorphism; genome mapping; physical mapping; fingerprinting;
KM      variety identification; genetic variability evaluation; primer; ss.
XX
OS      Eucalyptus grandis.
XX
PN      MO9967421-A1.
XX
PD      29-DEC-1999.
XX
PF      25-JUN-1999; 99WO-NZ00092.
XX
PR      25-JUN-1998; 98US-0105307.
XX
PA      (GENE-) GENESIS RES & DEV CORP LTD & FLETCHER.
PA      (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI      Havukkala IJ, Bloksberg LN, Glenn M;
XX
DR      WPI; 2000-116958/10.
XX
PT      New plant microsatellite markers and associated flanking species for
PT      the detection of polymorphic genetic markers -
XX
PS      Claim 1; Page 181; 392pp; English.
XX
CC      Sequences AAA31040-A32093 represent novel plant microsatellite sequences
CC      and associated flanking species. The sequences comprise a central core
CC      repeat sequence, especially selected from the sequences AAA32094-A32096
CC      with left and right flanking sequences. The polynucleotide sequences
CC      can be used in the detection of DNA polymorphisms, in genome mapping,
CC      in physical mapping, in positional cloning of genes, in variety
CC      identification and in evaluation of genetic variability within and
CC      between plant tissues, populations, cultivars, species and species
CC      groups. They may also be used to design hybridization probes for
CC      oligonucleotide fingerprinting and library screening and to design
CC      primers for microsatellite-primed PCR. Microsatellite markers are
CC      useful to locate specific economically useful genes in plant genomes.
XX
SO      Sequence 367 BP; 76 A; 116 C; 103 G; 72 T; 0 other;

Alignment Scores:
Pred. No.:          6.45          Length:          367
Score:              9.00          Matches:          9
Percent Similarity: 100.00%      Conservative:    0
Best Local Similarity: 100.00%    Mismatches:     0
Query Match:        1.81%        Indels:          0
DB:                 21           Gaps:           0

US-09-872-523-5 (1-498) x AAA31392 (1-367)
OY      179 LysValGlyGlnArgLeuGluLeu 187
DB      256 AAGGTAGACAGCGTCTCGAACTTCTT 230

RESULT 13
AAA31276/c
ID      AAA31276 standard; DNA: 383 BP.
XX

```

```

AC      AAA31276;
XX
DT      05-JUL-2000 (first entry)
XX
DE      Plant microsatellite marker #237.
XX
KM      Plant microsatellite sequence; core repeat sequence; detection; probe;
KM      DNA polymorphism; genome mapping; physical mapping; fingerprinting;
KM      variety identification; genetic variability evaluation; primer; ss.
XX
OS      Eucalyptus grandis.
XX
PN      MO9967421-A1.
XX
PD      29-DEC-1999.
XX
PF      25-JUN-1999; 99WO-NZ00092.
XX
PR      25-JUN-1998; 98US-0105307.
XX
PA      (GENE-) GENESIS RES & DEV CORP LTD & FLETCHER.
PA      (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI      Havukkala IJ, Bloksberg LN, Glenn M;
XX
DR      WPI; 2000-116958/10.
XX
PT      New plant microsatellite markers and associated flanking species for
PT      the detection of polymorphic genetic markers -
XX
PS      Claim 1; Page 145; 392pp; English.
XX
CC      Sequences AAA31040-A32093 represent novel plant microsatellite sequences
CC      and associated flanking species. The sequences comprise a central core
CC      repeat sequence, especially selected from the sequences AAA32094-A32096
CC      with left and right flanking sequences. The polynucleotide sequences
CC      can be used in the detection of DNA polymorphisms, in genome mapping,
CC      in physical mapping, in positional cloning of genes, in variety
CC      identification and in evaluation of genetic variability within and
CC      between plant tissues, populations, cultivars, species and species
CC      groups. They may also be used to design hybridization probes for
CC      oligonucleotide fingerprinting and library screening and to design
CC      primers for microsatellite-primed PCR. Microsatellite markers are
CC      useful to locate specific economically useful genes in plant genomes.
XX
SO      Sequence 383 BP; 84 A; 120 C; 105 G; 74 T; 0 other;

Alignment Scores:
Pred. No.:          6.72          Length:          383
Score:              9.00          Matches:          9
Percent Similarity: 100.00%      Conservative:    0
Best Local Similarity: 100.00%    Mismatches:     0
Query Match:        1.81%        Indels:          0
DB:                 21           Gaps:           0

US-09-872-523-5 (1-498) x AAA31276 (1-383)
OY      179 LysValGlyGlnArgLeuGluLeu 187
DB      270 AAGGTAGACAGCGTCTCGAACTTCTT 244

RESULT 14
AAA31391/c
ID      AAA31391 standard; DNA: 401 BP.
XX
AC      AAA31391;
XX
DT      05-JUL-2000 (first entry)
XX
DE      Plant microsatellite marker #352.
XX
KM      Plant microsatellite sequence; core repeat sequence; detection; probe;
KM      DNA polymorphism; genome mapping; physical mapping; fingerprinting;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: September 2, 2003, 20:33:10 : Search time 94 Seconds
(without alignments)
2338.391 Million cell updates/sec

Title: US-09-872-523-5

Perfect score: 498

Sequence: 1 MSELKIVRANKSDRLDK.....PIGCEAHSYVLPPKKYNY 498

Scoring table:

OLIGO
Xgapop 60.0 , Ygapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1135299

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=x1h
-O=/cgn2_1/USPRO.spool/US09872523/runat_02092003_073003_27299/app.query.fasta.1.647
-DB=Issued_Patents_NA -OPMT=fastap -SUFFIX=oli.rni -MINMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09872523.ecgn.1.1.56.ernat.02092003_073003_27299 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

1: Issued_Patents_NA:*
2: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	8	1.6	3526	1 US-08-323-084A-7	Sequence 7, Appl1
2	8	1.6	3526	1 US-08-323-084A-8	Sequence 8, Appl1
3	8	1.6	3526	1 US-08-674-008-7	Sequence 7, Appl1
4	8	1.6	3526	1 US-08-674-008-8	Sequence 8, Appl1
5	8	1.6	3986	4 US-09-053-374A-3	Sequence 3, Appl1
6	7	1.4	40	5 PCT-US94-06079-26	Sequence 26, Appl1
7	7	1.4	42	5 PCT-US94-06079-6	Sequence 6, Appl1
8	7	1.4	48	1 US-07-834-539A-64	Sequence 64, Appl1
9	7	1.4	48	1 US-08-053-131-112	Sequence 112, App
10	7	1.4	48	1 US-08-645-641-112	Sequence 112, App
11	7	1.4	48	1 US-07-853-408B-112	Sequence 112, App
12	7	1.4	48	1 US-08-096-762-112	Sequence 112, App

13	7	1.4	48	2 US-08-800-353-64	Sequence 64, Appl1
14	7	1.4	48	2 US-08-308-865-112	Sequence 112, App
15	7	1.4	48	3 US-09-042-353-305	Sequence 305, App
16	7	1.4	48	3 US-08-758-417A-153	Sequence 153, App
17	7	1.4	48	5 PCT-US92-06185-64	Sequence 64, Appl1
18	7	1.4	48	5 PCT-US92-10983-112	Sequence 112, App
19	7	1.4	87	4 US-09-465-901-5	Sequence 5, Appl1
20	7	1.4	119	1 US-07-634-278-80	Sequence 80, Appl1
21	7	1.4	119	1 US-08-477-728-80	Sequence 80, Appl1
22	7	1.4	119	1 US-08-474-040-80	Sequence 80, Appl1
23	7	1.4	119	1 US-08-487-200-80	Sequence 80, Appl1
24	7	1.4	119	3 US-08-484-537-80	Sequence 80, Appl1
25	7	1.4	146	1 US-08-116-388-9	Sequence 9, Appl1
26	7	1.4	146	5 PCT-US94-09318-9	Sequence 9, Appl1
27	7	1.4	162	4 US-08-746-411A-3	Sequence 3, Appl1
28	7	1.4	162	4 US-08-857-046A-3	Sequence 3, Appl1
29	7	1.4	162	4 US-09-573-252-3	Sequence 3, Appl1
30	7	1.4	248	5 PCT-US94-06079-8	Sequence 8, Appl1
31	7	1.4	273	4 US-09-313-294A-4219	Sequence 4219, App
32	7	1.4	275	4 US-09-313-294A-2176	Sequence 2176, App
33	7	1.4	282	4 US-09-216-393B-48	Sequence 48, Appl1
34	7	1.4	297	4 US-09-313-294A-2902	Sequence 2902, App
35	7	1.4	312	4 US-08-633-109-10	Sequence 10, Appl1
36	7	1.4	318	4 US-08-635-109-11	Sequence 11, Appl1
37	7	1.4	318	4 US-08-635-109-11	Sequence 11, Appl1
38	7	1.4	321	2 US-08-378-939-13	Sequence 13, Appl1
39	7	1.4	321	3 US-08-838-682-17	Sequence 17, Appl1
40	7	1.4	321	3 US-08-838-682-18	Sequence 18, Appl1
41	7	1.4	321	3 US-08-895-914-17	Sequence 17, Appl1
42	7	1.4	321	3 US-08-895-914-18	Sequence 18, Appl1
43	7	1.4	321	3 US-09-240-274-197	Sequence 197, App
44	7	1.4	321	3 US-09-357-710A-17	Sequence 17, Appl1
45	7	1.4	321	3 US-09-357-710A-18	Sequence 18, Appl1

ALIGNMENTS

RESULT 1
US-08-323-084A-7
Sequence 7, Application US/08323084A
Patent No. 5563048
GENERAL INFORMATION:
APPLICANT: HONJO, TASUKU
APPLICANT: SHIROZU, MICHIO
APPLICANT: TADA, HIDEAKI
TITLE OF INVENTION: No. 5563048el Polypeptides and DNAs encoding them
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRUE, MIOW, ZINN, MACPHEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/323.084A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 280505/1993
FILING DATE: 14-OCT-1993
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)293-7060
TELEFAX: (202)293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3526 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-323-084A-7

Alignment Scores:

Pred. No.:	128	Length:	3526
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.61%	Indels:	0
DB:	1	Gaps:	0

US-09-872-523-5 (1-498) x US-08-323-084A-7 (1-3526)

QY 352 GluAspSerPheProIleHisIle 359
DB 1563 GAAGATCTTCCCAATTCACATC 1586

RESULT 2

US-08-323-084A-8
Sequence 8, Application US/08323084A

Patent No. 5563048

GENERAL INFORMATION:

APPLICANT: HONJO, TASUKU

APPLICANT: SHIROZU, MICHIO

APPLICANT: TADA, HIDEAKI

TITLE OF INVENTION: No. 5563048el Polypeptides and DNAs encoding them

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS

STREET: 2100 Pennsylvania Avenue, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20037-3202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/323,084A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 280505/1993

FILING DATE: 14-OCT-1993

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)293-7060

TELEFAX: (202)293-7860

TELEX: 6491103

INFORMATION FOR SEQ. ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 3526 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

CELL LINE: FLEB14

FEATURE:

NAME/KEY: CDS

LOCATION: 80..361

IDENTIFICATION METHOD: by similarity to some other pattern

FEATURE:

NAME/KEY: sig_peptide

LOCATION: 80..142

IDENTIFICATION METHOD: by similarity with known sequence or

IDENTIFICATION METHOD: to an established consensus

NAME/KEY: mat_peptide
LOCATION: 143..358
IDENTIFICATION METHOD: by similarity with known sequence or
IDENTIFICATION METHOD: to an established consensus
US-08-323-084A-8

Alignment Scores:

Pred. No.:	128	Length:	3526
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.61%	Indels:	0
DB:	1	Gaps:	0

US-09-872-523-5 (1-498) x US-08-323-084A-8 (1-3526)

QY 352 GluAspSerPheProIleHisIle 359
DB 1563 GAAGATCTTCCCAATTCACATC 1586

RESULT 3

US-08-674-008-7
Sequence 7, Application US/08674008

Patent No. 5756084

GENERAL INFORMATION:

APPLICANT: HONJO, Tasuku

APPLICANT: SHIROZU, Michio

APPLICANT: TADA, Hideaki

TITLE OF INVENTION: HUMAN STROMAL DERIVED

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS

STREET: 2100 Pennsylvania Avenue, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20037-3202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/674,008

FILING DATE: 1-JUL-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/323,084

FILING DATE: 14-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 280505/1993

FILING DATE: 14-OCT-1993

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)293-7060

TELEFAX: (202)293-7860

TELEX: 6491103

INFORMATION FOR SEQ. ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 3526 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

US-08-674-008-7

Alignment Scores:

Pred. No.:	128	Length:	3526
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.61%	Indels:	0
DB:	1	Gaps:	0

US-09-872-523-5 (1-498) x US-08-674-008-7 (1-3526)

OY 352 GluAspSerPheProIleHisIle 359

DB 1563 GAAGATTCTTCCCAATTCACATC 1586

RESULT 4

US-08-674-008-8

; Sequence 8, Application US/08674008

; Patent No. 5736084

; GENERAL INFORMATION:

; APPLICANT: HONJO, Tasuku

; APPLICANT: SHIROZU, Michio

; APPLICANT: TADA, Hideaki

; TITLE OF INVENTION: HUMAN STROMAL DERIVED

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SUGHRUE, MION, ZINN, MACPHEAR & SEAS

; STREET: 2100 Pennsylvania Avenue, N.W.

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20037-3202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; FILING DATE: 1-JUL-1996

; CLASSIFICATION: 435

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 08/323,084

; FILING DATE: 14-OCT-1994

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: JP 280505/1993

; FILING DATE: 14-OCT-1993

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)293-7060

; TELEFAX: (202)293-7860

; TELEX: 6491103

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3526 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA to mRNA

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

; CELL LINE: FLEB14

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 80..361

; IDENTIFICATION METHOD: by similarity to some other pattern

; FEATURE:

; NAME/KEY: sig-peptide

; LOCATION: 80..142

; IDENTIFICATION METHOD: by similarity with known sequence

; FEATURE:

; NAME/KEY: mat-peptide

; LOCATION: 143..358

; IDENTIFICATION METHOD: by similarity with known sequence

; IDENTIFICATION METHOD: or to an established consensus

US-08-674-008-8

Alignment Scores:

Pred. No.: 128

Score: 8.00

Length: 3526

Matches: 8

Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.61%
DB: 1
Gaps: 0

US-09-872-523-5 (1-498) x US-08-674-008-8 (1-3526)

OY 352 GluAspSerPheProIleHisIle 359

DB 1563 GAAGATTCTTCCCAATTCACATC 1586

RESULT 5

US-09-053-374A-3/C

; Sequence 3, Application US/09053374A

; Patent No. 6462177

; GENERAL INFORMATION:

; APPLICANT: YEN, KWANG-MU

; TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: AMGEN INC.

; STREET: ONE AMGEN CENTER DRIVE

; CITY: THOUSAND OAKS

; STATE: CA

; COUNTRY: US

; ZIP: 91320

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/053,374A

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: COOK, ROBERT R.

; REGISTRATION NUMBER: 31,602

; REFERENCE/DOCKET NUMBER: A-514

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3986 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-09-053-374A-3

Alignment Scores:

Pred. No.: 144

Score: 8.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 1.61%

DB: 4

Gaps: 0

US-09-872-523-5 (1-498) x US-09-053-374A-3 (1-3986)

OY 426 ArgValGlyLeuArgLeuGlnAla 433

DB 375 AGGGTCGGGCTGAGCTGAGGCG 352

RESULT 6

PCT-US94-06079-26/C

; Sequence 26, Application PC/TUS9406079

; GENERAL INFORMATION:

; APPLICANT: American Cyanamid Company, .

; TITLE OF INVENTION: Gene Insertion by Direct Ligation In

; NUMBER OF SEQUENCES: 45

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: American Cyanamid Company

; STREET: One Cyanamid Plaza

CITY: Wayne
STATE: New Jersey
COUNTRY: USA
ZIP: 07470-8426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/06079
FILING DATE: 27-MAY-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Gordon, Alan M.
REGISTRATION NUMBER: 30,637
REFERENCE/DOCKET NUMBER: 31969-00\PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-831-3244
TELEFAX: 201-831-3305
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
PCT-US94-06079-26

Alignment Scores:
Pred. No.: 18.9 Length: 40
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.41% Indels: 0
DB: 5 Gaps: 0

US-09-872-523-5 (1-498) x PCT-US94-06079-26 (1-40)

QY 428 GlyLeuArgLeuGluAlaAla 434
|||||
DB 21 GGCGTGAGGCTCGAGGCTGCT 1

RESULT 7
PCT-US94-06079-6/c
Sequence 6, Application PC/TUS9406079
GENERAL INFORMATION:
APPLICANT: American Cyanamid Company,
TITLE OF INVENTION: Gene Insertion by Direct Ligation In
VITRO
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: New Jersey
COUNTRY: USA
ZIP: 07470-8426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/06079
FILING DATE: 27-MAY-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Gordon, Alan M.
REGISTRATION NUMBER: 30,637
REFERENCE/DOCKET NUMBER: 31969-00\PCT

TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-831-3244
TELEFAX: 201-831-3305
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
PCT-US94-06079-6

Alignment Scores:
Pred. No.: 19.8 Length: 42
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.41% Indels: 0
DB: 5 Gaps: 0

US-09-872-523-5 (1-498) x PCT-US94-06079-6 (1-42)

QY 428 GlyLeuArgLeuGluAlaAla 434
|||||
DB 21 GGCGTGAGGCTCGAGGCTGCT 1

RESULT 8
US-07-834-539A-64
Sequence 64, Application US/07834539A
GENERAL INFORMATION:
APPLICANT: Lonberg, Nils
APPLICANT: Kay, Robert M.
TITLE OF INVENTION: Transgenic Non-Human Animals Capable of
Producing Heterologous Antibodies
NUMBER OF SEQUENCES: 77
CORRESPONDENCE ADDRESS:
ADDRESSEE: William M. Smith
STREET: One Market Plaza, Steuart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/834,539A
FILING DATE: 1992-02-05
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 14643-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-543-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-07-834-539A-64

Alignment Scores:
Pred. No.: 22.6 Length: 48
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.41% Indels: 0
DB: 1 Gaps: 0

US-09-872-523-5 (1-498) x US-07-834-539A-64 (1-48)

QY 158 GYGLnThrSerLeuSerPro 164

DB 23 GGACAGACTCTCACCACCA 43

RESULT 9

US-08-053-131-112

; Sequence 112, Application US/08053131

; Patent No. 5661016

; GENERAL INFORMATION:

; APPLICANT: Lonberg, Nils

; TITLE OF INVENTION: Transgenic No. 5661016-Human Animals for

; TITLE OF INVENTION: Producing Heterologous Antibodies

; NUMBER OF SEQUENCES: 197

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend Hourie and Crew

; STREET: One Market Plaza, Steuart Tower, Suite 200

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94105

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/053,131

; FILING DATE: 26-Apr-1993

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/990,860

; FILING DATE: 16-DEC-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/810,279

; FILING DATE: 17-DEC-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/853,408

; FILING DATE: 18-MAR-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Smith, William M.

; REGISTRATION NUMBER: 30,223

; REFERENCE/DOCKET NUMBER: 14643-9-3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-326-2400

; TELEFAX: 415-326-2422

; INFORMATION FOR SEQ ID NO: 112:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 48 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (primer)

; US-08-053-131-112

Alignment Scores:

Pred. No.: 22.6 Length: 48

Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.41% Indels: 0

DB: 1 Gaps: 0

US-09-872-523-5 (1-498) x US-08-053-131-112 (1-48)

QY 158 GYGLnThrSerLeuSerPro 164

DB 23 GGACAGACTCTCACCACCA 43

DB 23 GGACAGACTCTCACCACCA 43

RESULT 10

US-08-645-641-112

; Sequence 112, Application US/08645641

; Patent No. 5719032

; GENERAL INFORMATION:

; APPLICANT: Lonberg, Nils

; TITLE OF INVENTION: Transgenic No. 5719032-Human Animals for

; TITLE OF INVENTION: Producing Heterologous Antibodies

; NUMBER OF SEQUENCES: 150

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: William M. Smith

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/645,641

; FILING DATE: 20-MAY-1996

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/904,068

; FILING DATE: 23-JUN-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Smith, William M.

; REGISTRATION NUMBER: 30,223

; REFERENCE/DOCKET NUMBER: 14643-000913

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-326-2400

; TELEFAX: 415-326-2422

; INFORMATION FOR SEQ ID NO: 112:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 48 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (primer)

; US-08-645-641-112

Alignment Scores:

Pred. No.: 22.6 Length: 48

Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.41% Indels: 0

DB: 1 Gaps: 0

US-09-872-523-5 (1-498) x US-08-645-641-112 (1-48)

QY 158 GYGLnThrSerLeuSerPro 164

DB 23 GGACAGACTCTCACCACCA 43

RESULT 11

US-07-853-408B-112

; Sequence 112, Application US/07853408B

; Patent No. 5789650

; GENERAL INFORMATION:

; APPLICANT: Lonberg, Nils

; TITLE OF INVENTION: Transgenic No. 5789650-Human Animals for

; TITLE OF INVENTION: Producing Heterologous Antibodies

; NUMBER OF SEQUENCES: 150

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: William M. Smith

STREET: One Market Plaza, Stewart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/853,408B
FILING DATE: 19920318
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 14643-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 112:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (primer)
US-07-853-408B-112

Alignment Scores:
Pred. No.: 22.6 Length: 48
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.41% Indels: 0
DB: 1 Gaps: 0

US-09-872-523-5 (1-498) x US-07-853-408B-112 (1-48)
QY 158 GYGlnThrSerLeuSerPro 164
DB 23 GGACGACTCTCCTCACA 43

RESULT 12
US-08-096-762-112
Sequence 112, Application US/08096762
Patent No. 5814318
GENERAL INFORMATION:
APPLICANT: Lonberg, Nils
APPLICANT: Kay, Robert M.
TITLE OF INVENTION: Transgenic No. 5814318-Human Animals for
TITLE OF INVENTION: Producing Heterologous Antibodies
NUMBER OF SEQUENCES: 210
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: One Market Plaza, Stewart Tower, Suite 200
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/096,762
FILING DATE: 22-JUL-1993
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/053,131
FILING DATE: 26-APR-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/990,860
FILING DATE: 16-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/904,068
FILING DATE: 23-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/853,408
FILING DATE: 18-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/810,279
FILING DATE: 17-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 14643-9-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 112:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (primer)
US-08-096-762-112

Alignment Scores:
Pred. No.: 22.6 Length: 48
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.41% Indels: 0
DB: 1 Gaps: 0

US-09-872-523-5 (1-498) x US-08-096-762-112 (1-48)
QY 158 GYGlnThrSerLeuSerPro 164
DB 23 GGACGACTCTCCTCACA 43

RESULT 13
US-08-800-353-64
Sequence 64, Application US/08800353
Patent No. 5874299
GENERAL INFORMATION:
APPLICANT: Lonberg, Nils
APPLICANT: Kay, Robert M.
TITLE OF INVENTION: Transgenic No. 5874299-Human Animals Capable of
TITLE OF INVENTION: Producing Heterologous Antibodies
NUMBER OF SEQUENCES: 77
CORRESPONDENCE ADDRESS:
ADDRESSEE: William M. Smith
STREET: One Market Plaza, Stewart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/800,353
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/834,539
FILING DATE: 1992-02-05
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.

REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 14643-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-543-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-800-353-64

Alignment Scores:
Pred. No.: 22.6 Length: 48
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.41% Indels: 0
DB: 2 Gaps: 0

US-09-872-523-5 (1-498) x US-08-800-353-64 (1-48)

QY 158 G1yGlnThrSerLeuSerPro 164
|||||

Db 23 GGACAGACTTCACTCTCACCA 43

RESULT 14
US-08-308-865-112
Sequence 112, Application US/08308865
Patent No. 5877397
GENERAL INFORMATION:
APPLICANT: Lonberg, Nils
TITLE OF INVENTION: Transgenic No. 5877397-Human Animals for
TITLE OF INVENTION: Producing Heterologous Antibodies
NUMBER OF SEQUENCES: 150
CORRESPONDENCE ADDRESS:
ADDRESSEE: William M. Smith
STREET: One Market Plaza, Stewart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/308,865
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/145,707
FILING DATE:
APPLICATION NUMBER: US 07/904,068
FILING DATE: 23-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 14643-9-1-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 112:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (primer)

US-08-308-865-112

Alignment Scores:
Pred. No.: 22.6 Length: 48
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.41% Indels: 0
DB: 2 Gaps: 0

US-09-872-523-5 (1-498) x US-08-308-865-112 (1-48)

QY 158 G1yGlnThrSerLeuSerPro 164
|||||

Db 23 GGACAGACTTCACTCTCACCA 43

RESULT 15
US-09-042-353-305
Sequence 305, Application US/09042353
Patent No. 6255458
GENERAL INFORMATION:
APPLICANT: Lonberg, Nils
TITLE OF INVENTION: Transgenic No. 6255458-Human Animals for
TITLE OF INVENTION: Producing Heterologous Antibodies
NUMBER OF SEQUENCES: 421
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,353
FILING DATE: 13-MAR-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/810,279
FILING DATE: 17-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/853,408
FILING DATE: 18-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/904,068
FILING DATE: 23-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/990,860
FILING DATE: 16-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/053,131
FILING DATE: 26-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/096,762
FILING DATE: 22-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/155,301
FILING DATE: 18-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,739
FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/165,699
FILING DATE: 10-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/209,741
FILING DATE: 09-MAR-1994
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/352,322
FILING DATE: 07-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/544,404
FILING DATE: 10-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/728,463
FILING DATE: 10-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US96/16433
FILING DATE: 10-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/758,417
FILING DATE: 02-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/21803
FILING DATE: 01-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 014643-009040US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 305:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-09-042-353-305

Alignment Scores:

Pred. No.:	22.6	Length:	48
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.41%	Indels:	0
DB:	3	Gaps:	0

US-09-872-523-5 (1-498) x US-09-042-353-305 (1-48)

QY 158 GlynhrSerleuserPro 164
DB 23 GGACGACTCTCCTCACA 43

Search completed: September 2, 2003, 23:12:18
Job time : 107 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 2, 2003, 20:40:10 ; Search time 369 seconds
(without alignments)
3096.306 Million cell updates/sec

Title: US-09-872-523-5

Perfect score: 498
Sequence: 1 MSEFLKIVRANKRSDRKLDK.....PIGWCEAHNSVYLQPKKKNY 498

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1533700 segs, 1147125425 residues

Word size: 1

Total number of hits satisfying chosen parameters: 3063161

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US09872523/runat_02092003_073004_27329/app_query.fasta_1.647
-DB=Published.Applications_MA -QFMT=fastcap -SUFFIX=oli.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oliigo
-TRANS=human4.0.coi -LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09872523.@CGN.1.1.244 @runat_02092003_073004_27329
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FCGAPOP=6 -FCGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database: Published.Applications_MA:*

1: /cgn2.6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2.6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2.6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2.6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2.6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2.6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2.6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2.6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
9: /cgn2.6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2.6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2.6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2.6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2.6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2.6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2.6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16: /cgn2.6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2.6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	498	100.0	1497	10	US-09-872-523-6 Sequence 6, Appl1

2	411	82.5	1497	10	US-09-872-523-73	Sequence 73, Appl
3	411	79.7	1497	10	US-09-872-523-74	Sequence 74, Appl
4	397	79.7	1497	10	US-09-872-523-75	Sequence 75, Appl
5	278	55.8	2307	10	US-09-872-523-76	Sequence 76, Appl
6	278	55.8	2307	10	US-09-872-523-77	Sequence 77, Appl
7	278	55.8	2307	10	US-09-872-523-78	Sequence 78, Appl
8	234	1.6	234	10	US-09-878-574-13281	Sequence 13281, A
9	437	1.6	437	13	US-10-027-632-19030	Sequence 19030, A
10	590	1.6	590	13	US-10-027-632-226880	Sequence 226880,
11	590	1.6	590	13	US-10-027-632-226881	Sequence 226881,
12	608	1.6	608	13	US-10-027-632-222125	Sequence 222125,
13	609	1.6	609	13	US-10-027-632-6153	Sequence 6153, Ap
14	611	1.6	611	13	US-10-027-632-224608	Sequence 224608,
15	611	1.6	611	13	US-10-027-632-224609	Sequence 224609,
16	632	1.6	632	13	US-10-027-632-141937	Sequence 141937,
17	669	1.6	669	13	US-10-027-632-96511	Sequence 96511, A
18	669	1.6	669	13	US-10-027-632-306237	Sequence 306237,
19	833	1.6	833	13	US-10-027-632-166547	Sequence 166547,
20	833	1.6	833	13	US-10-144-929-63	Sequence 63, Appl
21	3524	1.6	3524	8	US-08-927-939-29	Sequence 29, Appl
22	3524	1.6	3524	10	US-09-919-497-45	Sequence 45, Appl
23	3524	1.6	3524	10	US-09-880-107-2204	Sequence 2204, Ap
24	4313	1.6	4313	10	US-09-070-927A-67	Sequence 67, Appl
25	9424	1.6	9424	8	US-08-910-386A-8	Sequence 8, Appl1
26	19011	1.6	19011	11	US-09-764-891-5813	Sequence 5813, Ap
27	19011	1.6	19011	11	US-09-764-891-5814	Sequence 5814, Ap
28	1.6	271990	14	US-10-195-144-87	Sequence 87, Appl	
29	1.6	1601042	13	US-10-027-632-55064	Sequence 59064, A	
30	87	1.4	87	14	US-10-259-430-5	Sequence 5, Appl1
31	201	1.4	201	13	US-10-027-632-24317	Sequence 24317, A
32	210	1.4	210	9	US-09-764-869-313	Sequence 313, Ap
33	210	1.4	210	14	US-10-091-504-313	Sequence 313, Ap
34	258	1.4	258	9	US-09-294-093B-781	Sequence 781, Ap
35	263	1.4	263	9	US-09-294-093B-3099	Sequence 3099, Ap
36	265	1.4	265	10	US-09-878-574-5748	Sequence 5748, Ap
37	269	1.4	269	10	US-09-796-692-3156	Sequence 3156, Ap
38	269	1.4	269	14	US-10-040-862-3156	Sequence 3156, Ap
39	275	1.4	275	9	US-09-923-876-5298	Sequence 5298, Ap
40	280	1.4	280	9	US-09-864-761-31987	Sequence 31987, A
41	280	1.4	280	10	US-09-878-574-15142	Sequence 15142, A
42	282	1.4	282	9	US-09-216-393-48	Sequence 48, Appl
43	285	1.4	285	10	US-09-878-574-5400	Sequence 5400, Ap
44	285	1.4	285	10	US-09-878-574-12434	Sequence 12434, A
45	285	1.4	285	14	US-10-102-524-1259	Sequence 1259, Ap

ALIGNMENTS

RESULT 1
US-09-872-523-6
Sequence 6, Application US/09872523
Patent No. US20020137906A1
GENERAL INFORMATION:
APPLICANT: Horvitz, H. Robert
APPLICANT: Davison, Ewa M.
APPLICANT: Lu, Xiaowei
TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
FILE REFERENCE: 01997/536002
CURRENT APPLICATION NUMBER: US/09/872,523
CURRENT FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: US 60/208,802
PRIOR FILING DATE: 2000-06-02
NUMBER OF SEQ ID NOS: 78
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 1497
TYPE: DNA
ORGANISM: Caenorhabditis elegans
US-09-872-523-6
Alignment Scores: 0
Pred. No.: 498.00
Length: 1497
Matches: 498

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-09-872-523-5 (1-498) x US-09-872-523-6 (1-1497)
QY 1 MetSerGluPheLeuLysIleValArgAlaAsnLysLysSerAspArgLysLeuAspLys 20
DB 1 ATGCTGAATTTCTGAAAAATTTGTCAGACCTTAACAAAAATCGCAGAAAACTCGATTAAG 60
QY 21 ThrTyrLeuTrpGluSerTyrLeuHisGlnPheGluLysGlyLysThrSerPheIlePro 40
DB 61 ACCTACTGTGGGAATCTATTATTCATCAGTTCCGAAAGGAAAAAATCTTTTCATTTCCA 120
QY 41 ValGluAlaPheAsnArgAsnLeuThrValAsnPheAsnGluCysValLysGluIleVal 60
DB 121 GTTGACAGATTCATTCGTAACCTTAACATTAATTTTAACGAATGGCTGAGAGAGAGATT 180
QY 61 IlePheGluTrpValValHisAspTyrAspLysAsnCysAspSerIleGlnValArgTrp 80
DB 181 ATCTTCGAACAGAGGTCATGATGATGACAAGAACTCGATTCGATTCAGATCAGATGG 240
QY 81 PheAlaArgIleGluLysValCysGlyTyrArgValIleuAlaGlnPheIleGlyAlaAsp 100
DB 241 TTTCGACGAATTGAAAAAGTTTGGGATACAGAGTTCTGGCTCAGATTATCGGAGCTGAC 300
QY 101 ThrLysPheTrpLeuAsnIleLeuSerAspMetPheGlyLeuAlaAsnAlaIleMet 120
DB 301 ACGAAATTTTGGCCCAATATTTTATCGACAGATGTTGGTTGGCAACGCCCAATG 360
QY 121 SerAspProAsnMetAspLysIleValTyrAlaProPheLeuAlaIleAsnGluLys 140
DB 361 AGTGATCCCAATATGGATAAATGTATATGCTCCGCCCTTGCAATCAAGAAATATC 420
QY 141 GlnAsnAspMetValAsnTyrValAsnAsnCysIleAspGlyGluIleValGlyGlnThr 160
DB 421 CAAATGATATGTAATTAATTAATTAATTCATGATGGCGAAATGCTCGCCAAACT 480
QY 161 SerLeuSerProLysPheAspGluLysAlaLeuLeuSerLysHisArgPheLysVal 180
DB 481 TCGCTGCTCCAAATTCGATGAGGAGGAGGCTCTCTAAGCAACATCGTTTCAAGCT 540
QY 181 GlyGlnArgLeuGluLeuLeuAsnTyrSerAsnSerThrGluIleArgValAlaArgIle 200
DB 541 GGACCAACCTCTGGAACCTAATTAATTAATTCATTCCTGGAATTAACGGGTACCGGAAT 600
QY 201 GlnGluIleCysGlyArgArgMetAsnValSerIleThrLysLysAspPheProGluSer 220
DB 601 CAAGAAATATGTGGACGACGAGATGATATCTATCAAAAGAAAGACTTCCCAATCG 660
QY 221 LeuProAspAlaAspAspAspArgGlnValPheSerSerGlySerGlnTyrTrpIleAsp 240
DB 661 CTTCGAGATGCGAGATGACGACAGACAAAGCTTTAGCTCTGATCTCAATATATGATAG 720
QY 241 GluGlySerPhePheIlePheProValGlyPheAlaAlaValAsnGlyTyrGlnLeuAsn 260
DB 721 GAGGGAACCTCTTCATATTCCTGCTGTGATTTGAGAGAGTCAAATGATATCACTAAT 780
QY 261 AlaLysLysGluTyrIleGluHisThrAsnLysIleAlaGlnAlaIleLysAsnGlyGlu 280
DB 781 GCGAAAAAGGAATATATGAGACACCAATAAATTTGCTCAACGAATAAAAATGAGAGA 840
QY 281 AsnProArgTyrAspSerAspAspValThrPheAspGlnLeuAlaLysAspProIleAsp 300
DB 841 AATCCAAATATGACTCAGACGACCTCAATTTGATCAATTAAGCAAAAGATCCCAATGAT 900
QY 301 ProMetIleTrpArgLysValLysValGlyGlnLysPheGluLeuIleAspProLeuAla 320
DB 901 CCCATGATTTGGAGAAAGTTAAGGTGACAAAAAGTTGAGCTCATGACCCCTTGCT 960
QY 321 GlnGlnPheAsnAsnLeuHisValAlaSerIleLeuLysPheCysLysThrGluLysTyr 340

DB 961 CAGCAATTCATTAACCTCCAGCTGCTTCGATTCTCAAAATTTTGCAAAATCGAAGATAT 1020
QY 341 LeuIleValGlyMetAspGlyProAspAlaLeuGluAspSerPheProIleHisIleAsn 360
DB 1021 CTATATTTGGGAGATGATGATGCTCAGATGCTACCTTGATGAAGACAGTTTCCATATCATAT 1080
QY 361 AsnThrPheMetPheProValGlyTyrAlaGluLysTyrAsnLeuGluLeuValProPro 380
DB 1081 AATCAATTTATGTTCCCGAGTTGGTTATTCGGAAGAAATTAATTTTGGAACTGTTCCGCCA 1140
QY 381 AspGluPheLysGlyThrPheArgTrpAspGluTyrLeuGluLysGluSerAlaGluThr 400
DB 1141 GATGACTTCGAAGACACTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
QY 401 LeuProLeuAspLeuPheLysProMetProSerGlnGluArgLeuAspLysPheLysVal 420
DB 1201 CTACCGCTTGACTGTTTCAACCAATGCTTCCCAAGAGAGATTAGACAAATTTAAGSTA 1260
QY 421 IleLeuIleSerLysArgValGlyLeuArgLeuGluAlaAlaAspMetCysGluAsnGln 440
DB 1261 ATTCGATTTCCAAACGGGTAGGACTACGCTTGAAAGCTGCTGACATGTGAAATACAG 1320
QY 441 PheIleCysProAlaThrValLysSerValHisGlyArgLeuIleAsnValAsnPheAsp 460
DB 1321 TTTATTTTCCAGCTACAGTGAATCACTTATGATGATGATGATGATGATGATGATGAT 1380
QY 461 GlyTrpAspGluGluPheAspGluLeuTyrAspValAspSerHisAspIleLeuProIle 480
DB 1381 GGCTGGATGGAAGAAATTTGATGAACCTGTATGATGGTACCTCCATGATATTTCCCGATA 1440
QY 481 GlyTrpCysGluAlaHisSerTyrValLeuGlnProPheLysLysTyrAsnTyr 498
DB 1441 GGATGGTGTGAAGCGCACAGTTATGCTTACAACTCCGAAAAAGTACAACTAT 1494
RESULT 2
US-09-872-523-73
; Sequence 73, Application US/09872523
; Patent No. US20020137906A1
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Davidson, Ewa M.
; APPLICANT: Lu, Xiaowei
; TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
; FILE REFERENCE: 01997/536002
; CURRENT APPLICATION NUMBER: US/09/872,523
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/208,802
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 1497
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-872-523-73
Alignment Scores:
Pred. No.: 0 Length: 1497
Score: 411.00 Matches: 497
Percent Similarity: 99.60% Conservative: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 82.53% Indels: 2
DB: 10 Gaps: 0
US-09-872-523-5 (1-498) x US-09-872-523-73 (1-1497)
QY 1 MetSerGluPheLeuLysIleValArgAlaAsnLysLysSerAspArgLysLeuAspLys 20
DB 1 ATGCTGAATTTCTGAAAAATTTGTCAGACCTTAACAAAAATCGCAGAAAACTCGATTAAG 60
QY 21 ThrTyrLeuTrpGluSerTyrLeuHisGlnPheGluLysGlyLysThrSerPheIlePro 40
DB 61 ACCTACTGTGGGAATCTATTATTCATCAGTTCCGAAAGGAAAAAATCTTTTCATTTCCA 120

```

QY 41 ValGluAlaPheAsnArgAsnLeuThrValAsnPheAsnGluCysValIysGluGlyVal 60
   |||||
Db 121 GTTGAAGCATTCATGCTACCTTACAGTTAAATTTTAAACGAATGCTGGAAGAGAGATT 180
QY 61 IlePheGluThrValValHisAspTyrAspLysAsnCysAspSerIleGlnValArgTyr 80
   |||||
Db 181 ATCTTGAAACAGATGGTCCATGATTTATGACAAGAACTGGCAATTCGATTCAGTCAAGTGG 240
QY 81 PheAlaArgIleGluLysValCysGlyTyrArgValLeuAlaGlnPheIleGlyAlaAsp 100
   |||||
Db 241 TTTTGACGAAATTGAAAAAGTTTGGCGATACAGAGTTCTGCTCAGTTATCGCAGCTGAC 300
QY 101 ThrLysPheTyrLeuAsnIleLeuSerAspAspMetPheGlyLeuAlaAsnAlaIleMet 120
   |||||
Db 301 ACGAAATTTTGGCTCATATTTTATGCGACGATATGTTGGTTGGCAACGCCGCAAGT 360
QY 121 SerAspProAsnMetAspLysIleValTyrAlaProProLeuAlaIleAsnGluGlyTyr 140
   |||||
Db 361 AGTATCCCAATATGATGATAAATTTGATATGCTCCGCCGTTGCATTCACAGAAATAC 420
QY 141 GlnAsnAspMetValAsnTyrValAsnAsnCysIleAspGlyGluIleValGlyGlnTyr 160
   |||||
Db 421 CAAAAGATATGATAATTAATGTAATTAATTCATTGATGCGCAAAATCGTCCGCCAAAGT 480
QY 161 SerLeuSerProLysPheAspGluGlyLysAlaLeuLeuSerLysHisArgPheLysVal 180
   |||||
Db 481 TCGCTCTCCCAAAATTCATGACGAGGAAGGCTCTCTTAAGCAAGATCGTTCAAAAGTT 540
QY 181 GlyGlnArgLeuGluLeuLeuAsnTyrSerAsnSerThrGluIleArgValAlaArgIle 200
   |||||
Db 541 GGACAACAGCTCTGAACTATAATTAATTTCCAAATTTACTGAAATACGCCGAGCGCAATT 600
QY 201 GlnGluIleCysGlyArgArgMetAsnValSerIleThrLysLysAspPheProGluSer 220
   |||||
Db 601 CAAAGAAATATGTGGACGACGAAGATGATATCATCACAAAGAAACCTTCCCGCAATCG 660
QY 221 LeuProAspAlaAspAspAspArgGlnValPheSerSerGlySerGlnTyrTrpIleAsp 240
   |||||
Db 661 CTTCCGATGACGATGACGACAGCAAGCTTTAGCTGCGATCACAATATTCGATGAC 720
QY 241 GluGlySerPhePheIlePheProValGlyPheAlaAlaValAsnGlyTyrGlnLeuAsn 260
   |||||
Db 721 GAGGGAGCTTCTCATATTTCTCTGTTGATTTGGCAGCTCAATGATATCAACTAAT 780
QY 261 AlaLysLysGluTyrIleGluHisThrAsnLysIleAlaGlnAlaIleLysAsnGlyGlu 280
   |||||
Db 781 GCGAAAGGAATATATTTAGCACACAAATTAATTTGCTCAACCAATTAATAATGAGAA 840
QY 281 AsnProArgTyrAspSerAspAspValThrPheAspGluLeuAlaLysAspProIleAsp 300
   |||||
Db 841 AATCCAAATATGACATCACACGACGTCACATTTGATTCATTAACAAAGATTCATGAT 900
QY 301 PrometIleTyrArgLysValLysValGlyGlnLysPheGluLeuIleAspProLeuAla 320
   |||||
Db 901 CCCATGATTTGGGAAAGATTAAGTTGAACAATAATTGAGCTCATACCAACCCCTTGCT 960
QY 321 GlnGlnPheAsnAsnLeuHisValAlaSerIleLeuLysPheCysLysThrGluGlyTyr 340
   |||||
Db 961 CAGCAATTCATTAACCTCCACGCTCGCTGATTCCTCAAAATTTGCAAACTGAAAGATAT 1020
QY 341 LeuIleValAlaGlyMetAspGlyProAspAlaLeuGluAspSerPheProIleHisIleAsn 360
   |||||
Db 1021 CTTATTTGTGGAAATGATGCTCAGATGACACTTGAAGACAGTTTCTTATTCATCAT 1080
QY 361 AsnThrPheMetPheProValGlyTyrAlaGluLysTyrAsnLeuGluLeuValProPro 380
   |||||
Db 1081 AATACATTTATGTCTCCACATTTGGTTATGCGGAAAGATTAATTTGGAACCTGTTCCGCCA 1140
QY 381 AspGluPheLysGlyThrPheArgTyrAspGluTyrLysGluLysGluSerAlaGluTyr 400
   |||||
Db 1141 GATGAGTTCAAGAGCAATTCAGATGGGATGATTAATCTTGAGAGAAATCTGCAGAAACC 1200

```

```

QY 401 LeuProLeuAspLeuPheLysPrometProSerGln-GluArgLeuAspLysPheLysVa 420
   |||||
Db 1201 CTACCGCTTGACTGTTGTAAGCAATGCCCTTCTTA-AGACAGATTTGACAAATTTAAGGT 1259
QY 420 IileLeuIleSerLysArgValGlyLeuArgLeuGluAlaAlaAspMetCysGluAsnGly 440
   |||||
Db 1260 AATTCGATTTCCAAACGGGTAGACTACCTTGAAGCTGTCGACATGTGTGTAATAATCA 1319
QY 440 nPheIleCysProAlaThrValLysSerValHisGlyArgLeuIleAsnValAsnPheAs 460
   |||||
Db 1320 GTTATTTTGCACGTACAGTGAATCAGTTCAATGGAAGCTGATTAATGTCAATTTCCA 1379
QY 460 pGlyTyrAspGluGlnPheAspGluLeuTyrAspValAspSerHisAspIleLeuProIle 480
   |||||
Db 1380 CGGCTGGAGTGAAGATTTGATGACTGATGATGATGACACTCCCATGATTTCTACCGAT 1439
QY 480 eGlyTyrCysGluAlaHisSerTyrValLeuGlnProProLysLysTyrAsnTyr 498
   |||||
Db 1440 AGGATGGTGGAGGCGACAGTATGTTCTACAACTCCGAAAGATCAACTAT 1494

```

RESULT 3

```

US-09-872-523-74
; Sequence 74, Application US/09872523
; Patent No. US20020137906a1
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Davison, Ewa M.
; APPLICANT: Lu, Xiaowei
; TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
; FILE REFERENCE: 0197/536002
; CURRENT APPLICATION NUMBER: US/09/872,523
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/208,802
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 1497
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-872-523-74

```

```

Alignment Scores:
Pred. No.: 0 Length: 1497
Score: 411.00 Matches: 497
Percent Similarity: 99.60% Conservative: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 82.53% Indels: 2
Gaps: 0

```

US-09-872-523-5 (1-498) x US-09-872-523-74 (1-1497)

```

QY 1 MetSerGluPheLeuLysIleValArgAlaAsnLysSerAspArgLysLeuAspLys 20
   |||||
Db 1 ATCTCGAAATTTCTGAAATTTGTCAGAGCTTAACAAAAATCGGACAGAAATCTCGATTAAG 60
QY 21 ThrTyrLeuTrpLysSerTyrGluHisGlnPheGluLysGlyLysThrSerPheIlePro 40
   |||||
Db 61 ACCTACTGTGGGAATCTATTTACATTCACCTTCGAGAAAGAAACCTTTTCATTCCA 120
QY 41 ValGluAlaPheAsnArgAsnLeuThrValAsnPheAsnGluCysValIysGluGlyVal 60
   |||||
Db 121 GTTGAAGCATTCATGATGTAACCTTACAGTTAAATTTTAAACGAATGCGGAAGAGAGATT 180
QY 61 IlePheGluThrValValHisAspTyrAspLysAsnCysAspSerIleGlnValArgTyr 80
   |||||
Db 181 ATCTTGAAACACATGTGCTCATGATTTATGACAAAGCTGCAATTCGATTCAGTCAAGATGG 240
QY 81 PheAlaArgIleGluLysValCysGlyTyrArgValLeuAlaGlnPheIleGlyAlaAsp 100
   |||||
Db 241 TTTTGACGAAATTAATAAAGTTTTCGGATACAGAGTTCTGCTGATGATTCGAGAGCTGAC 300
QY 101 ThrLysPheTyrLeuAsnIleLeuSerAspAspMetPheGlyLeuAlaAsnAlaIleMet 120

```

|||||
Db 301 ACGAAATTTTGGCTCAATATTTTATCGGACGATATGTTGGTGGCAAGCCGCAATG 360
OY SeraspproasmetasplysIIeValTyraIaProProleuAlaIIaasnIugIuTy 140
Db 361 AGTGTATCCCATATGATGAATAATGTATATGTCTCGCGCTTGCAATCAGAGAAATAC 420
OY 141 GlnasnasmetValasnTyraIasnasncysIIeaspgIugIuIleValIgluIthr 160
Db 421 CAAATGATATGTGAATTAATTAATGATTCATTTGATGGGAATATGTGGCCAAAT 480
OY SerIeuserpIolysPheaspgIugIyrsAlaIeIuseryshIasrgPheIysVal 180
Db 481 TCGGTGTCCAAATTCGATGAGGGAAGGCTCTCCTAAGCAACGATTCGTTCAAAGT 540
OY 181 GlyIuarIeugIuIeIuIeIusnTyrsSerIasSerThrIuIeIarValaIaIarIle 200
Db 541 GGACACAGCTTGAACATTAATTAATTCATTCATTCGAAATCGGGTAGCGGCAAT 600
OY 201 GlnGluIleCysGlyArGArGmetasnaValserIIethrIyIysAspPheProIuSer 220
Db 601 CAGAAATATGTGACGACGAGATGATCTATCTACACAAAGAAAGACTTCCCAATCG 660
OY 221 LeuproaspIaaspaspaspIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 240
Db 661 CTTCAGATGCGATGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 720
OY 241 GluGlySerPhePheIlePheProValGlyPheAlaIaIaIaIaIaIaIaIaIaIa 260
Db 721 GAGGAGACCTTCTCATATTTCCGTGTGATTCGACAGCAATGATATCACTAAT 780
OY 261 AlaIyIyIyGluTyriIleGluHisThrAsnIyIleAlaGlnAlaIIeIyAsnIyIu 280
Db 781 GCGAAAAAGATATATGATGACACCAATATAATTTGCTCAAGCAATATAAATATGGA 840
OY 281 AsnProArGlyrAspSerAspAspValThrPheaspgIuIeAlaIyAspProIleas 300
Db 841 AATCCAAATATGACTCAGACGACGACGACGACGACGACGACGACGACGACGACGAC 900
OY 301 ProMetIleTrrPaqIyIyValIyValIyIyIyIyIyIyIyIyIyIyIyIyIyIy 320
Db 901 CCCATGATTTGGAGAAAGATTAAAGCTTGACAAAAGTTGAGCTCATGACCCCTTGCT 960
OY 321 GlnGlnIlePheasnasIleuHisValaIaSerIleIeIuIyIyPheCysIyIyThgIuIy 340
Db 341 LeuIleValIyIyMetaspgIyProaspIaIeugIuIasPser-PheProIleHisIleas 360
OY 1021 CTTATGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1079
Db 360 nAsnThrPheMetPheProValGlyTyraIaGluIyIyIyIyIyIyIyIyIyIyIy 380
OY 1080 TAAATCAATTTATGTTCCCAAGTGGTTATGCGGAAAAAGTAAATTTGGAACTGTCCGC 1139
Db 380 oAspgIuIy 400
OY 1140 AGATAGITTCAAAGGAACTTCAGATGGATGATGATGATGATGATGATGATGATGAT 1199
Db 400 rIeupProIeasPleuPheIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 420
OY 1200 CCTACCGCTTGACTTGCTCAAGCCATGCTTCCCAAGAGAGATTAAGCAATTTAAGT 1259
Db 420 ILeIeIuIeIserIy 440
OY 1260 AATTCGATTTCCAAAGGGGTAGACTAGCTTGAAGCTGTGACATGTGGAATCA 1319
Db 440 nPheIleCysProAlaIthrValIyIySerValHisIyIyIyIyIyIyIyIyIyIy 460
OY 1320 GTTATATTTGCTCCAGCTACAGTAAATCAGTTTATGGAAGACATGAAATTAATTTGCA 1379
Db 460 pGIyTrpAspGluIyIyPheaspgIuIeIyIyIyIyIyIyIyIyIyIyIyIyIy 480

Db 1380 CGGCTGGGATGAGCAATTTGATGAACCTGATGATGTGACCTCCATGATATTTACCAT 1439
OY 480 eGIyTrpCysgluIaIaHisSerTyraIeugIuIeIuseryshIasrgPheIy 498
Db 1440 AGGATGGTGTGAAGGCGACAGTATGTTCTCAACCTCCGAAAAAGTACAACTAT 1494
RESULT 4
US-09-872-523-75
; Sequence 75, Application US/09872523
; Patent No. US20020137906A1
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Davison, Ewa M.
; TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
; FILE REFERENCE: 01997/536002
; CURRENT APPLICATION NUMBER: US/09/872,523
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/208,802
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 75
; LENGTH: 1497
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-872-523-75
Alignment Scores:
Pred. No.: 0 Length: 1497
Score: 397.00 Matches: 497
Percent Similarity: 99.60% Conservative: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 79.72% Indels: 2
DB: 10 Gaps: 0
US-09-872-523-5 (1-498) x US-09-872-523-75 (1-1497)
OY 1 MetSerGluPheIeIyIyIleValArGAlaIasnyIyIySerAspArGlyIeIusAsp 20
Db 1 ATGTGTGAATTTCTGAAAAATGTCAGAGCTAAACAAAAATCGACAGAAACTCGATAG 60
OY 21 ThrTyIeIuTrpGluSerTyriIeIuHisGlnPheGluIyIyIyIyIyIyIyIyIyIy 40
Db 61 ACCTACTGTGTGGAACTCTATTTACATGATTCGAGAAAGGAAAACTCTTTCATTTCA 120
OY 41 ValGluAlaPheasnaIyAsnIeIuThrValaIasPheasngIuIyIyIyIyIyIyIy 60
Db 121 GTTGAAGCATTCAACTGATACCTTACAGTTAATTTTAAAGAAATCGTGAAGAAAGAT 180
OY 61 ILePheGluIthrValaIaIaIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 80
Db 181 ATCTTCGAAACGTGTCTCATGATTAATGACAAAGACTCGATTCGATTCAGTCAATG 240
OY 81 PheAlaArGIIeGluIyIyValIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 100
Db 241 TTTGACGCAATTTGAAAGGTTTGGCGATACAGATTCGGCTCAATTTATGAGACTGAC 300
OY 101 ThrIyIyPheTrpIeIusnIleIeIuserysPheIyIyIyIyIyIyIyIyIyIyIy 120
Db 301 ACGAAATTTTGGCTCAATATTTTATCGGACGATGATGTTGGTGGCAAAACCCCAATG 360
OY 121 SeraspproasmetasplysIIeValTyraIaPro-ProIeuaIaIeasngIuIy 140
Db 361 AGTGAATCCCAATATGATTAATTTATATATGCTTC-GCCGCTTGAATCAACGAAGAATA 419
OY 140 rGlnasnasmetValasnTyraIasnasncysIIeaspgIyIuIleValIgluIthr 160
Db 420 CCAAAATGATATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 479
OY 160 rSerIeuserpIolysPheaspgIuIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 180
Db 480 TTCGCTGTCTCCAAATTCGATGAGGAGAGCTCTCCTAAGCAACGATCGTTCAAAGT 539

```

QY 180 lglglnargleuGluleuLeuasnTyrSerAsnSerThrGluIleArgValAlaArg11 200
    |||
Db 540 TGGACAACGCTTTGAACCTATTAAATTATTCCTCAATTTACTGAAATACGGGTACGGGAAT 599
QY 200 eglngluileCysGlyArgArgMetAsnValSerIleThrLysLysAspPheProGluSe 220
    |||
Db 600 TCAGAAATATGTGACGACGAAATGAATGATCTATCACAAGAAAGAACTTTCCCAATC 659
QY 220 rleuproaspAlaaspAspArgGlnValPheSerSerGlySerGlnTyrTrpIleAs 240
    |||
Db 660 GCTTCAGATGAGATGACGACGACAGACAGACACTTTTACTCTGTGATTCATATTGGATAGA 719
QY 240 pgluglySerPhePheIlePheProValGlyPheAlaAlaValAsnGlyTyrGlnLeuAs 260
    |||
Db 720 CGAGGAAACGCTTCTCATATTTCCTGTGATTTGACAGACAGTCAATGATACACTTAA 779
QY 260 naAlaLysLysGluTyrIleGluHisThrAsnLysIleAlaGlnAlaIleLysAsnGlyG1 280
    |||
Db 780 TCGAAAAAGAAATATATTGAGCACAACAATATAAATTGCTCAAGCAATATAAATGAGCA 839
QY 280 uasnproargTyrAspSerAspAspValThrPheAspGlnLeuAlaLysAspProIleAs 300
    |||
Db 840 AATCCAAAGATATGACTCAGACGACGTCACATTGTGATTAACAAGATCCAAATGTA 899
QY 300 pPromeIleTyrParGlyValLysValGlyGlnLysPheGluLeuIleAspProLeuAl 320
    |||
Db 900 TCCCATGATTTGGAGAAAAGTTAAGTTGACAAAAGTTTGACCTATGACCCCTTGGC 959
QY 320 aglnglnPheAsnAsnLeuHisValAlaSerIleLeuLysPheCysLysThrGluGlyTyr 340
    |||
Db 960 TCAGCAATTCATTAACCTCCAGCTCCCTCGATTCCTCAAAATTTTGCAAAAGTAAAGATA 1019
QY 340 rleuIleValGlyMetAspGlyProaspAlaLeuGlnLysAspSerPheProIleHisLys 360
    |||
Db 1020 TCTTATTTGGAGATGATGATGTCACATGACACTTGAAGACAGATTTCCATTCATATCAA 1079
QY 360 nasnThrPheMetPheProValGlyTyrAlaGlnLysTyrAsnLeuGlnLeuValAlaProp 380
    |||
Db 1080 TATACATTTATGTTCCAGTTGGTTATCGGAAAAGTTAATTTGCACTTCTCCGCC 1139
QY 380 oAspGluPheLysGlyThrPheArgTrpAspGluTyrLeuGlnLysGlnSerAlaGluTh 400
    |||
Db 1140 ACATGAGTTCAAAAGAACATTCAGATGGATGAAATCTTGAGAAAAGAAATCTCCAGAAAC 1199
QY 400 rleuproLeuaspLeuPheLysPheMetProSerGlnLysArgLeuAspLysPheLysVa 420
    |||
Db 1200 CCTACCGCTTGACTTGTCAAGCCAAATGCCCTTCCCAAGAGAGATTAAGACAAATTTAAGGT 1259
QY 420 lIleLeuIleSerLysArgValGlyLeuArgLeuGlnAlaAlaAspMetCysGlnAsnG1 440
    |||
Db 1260 AATTTCGATTTCCAAACGGGTAGAGACTACGCCCTTGAAAGCTGCGACATGTGTAAATCA 1319
QY 440 nPheIleCysProAlaThrValLysSerValHisGlyArgLeuIleAsnValAsnPheAs 460
    |||
Db 1320 GTTTATTTTCCAGCTACAGTGAATAATCAGTTCATGAAACACTGATTAATGTCAATTTCCA 1379
QY 460 pGlyTrpAspGlnGluPheAspGluLeuTyrAspValAspSerHisAspIleLeuProI1 480
    |||
Db 1380 CGGCTGGAGTGAAGAAATTTGATGAACTGATGATGTGAGACTCCCAAGATATTTTACCGAT 1439
QY 480 eglTyrCysGlnAlaHisSerTyrValLeuGlnPropLysLysTyrAsnTyr 498
    |||
Db 1440 AGAGTGTGTGAAGCCACAGTTATGTCTACAACTCCGAAAAAATTAACACTAT 1494

```

RESULT 5
US-09-872-523-76

; Sequence 76, Application US/09872523

; Patent No. US20020137906A1

; GENERAL INFORMATION:

; APPLICANT: Horvitz, H. Robert

; APPLICANT: Davison, Ewa M.

; APPLICANT: Lu, Xiaowei

```

; TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
; FILE REFERENCE: 01997/536002
; CURRENT APPLICATION NUMBER: US/09/872,523
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/208,802
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 2307
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-872-523-76

```

Alignment Scores:

Pred. No.:	2,84e-287	Length:	2307
Score:	278.00	Matches:	278
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	55.82%	Indels:	0
DB:	10	Gaps:	0

US-09-872-523-5 (1-498) x US-09-872-523-76 (1-2307)

```

QY 150 AsnCysIleAspGlyGluIleValGlyGlnThrSerLeuSerProLysPheAspGlyG1 169
    |||
Db 1138 AATTGCATTTGATGGCGAATCGTGGCGCAAACTTCGCTGCTCCAAATTCGATGAAGG 1197
QY 170 LysAlaLeuLeuSerLysHisArgPheLysValGlyGlnArgLeuGlnLeuAsnTyr 189
    |||
Db 1198 AAGGCTCTCTTAAGCAAGCATCTCTTCAAAAGTTGGACACAGCTTGAACATTAATTAAT 1257
QY 190 SerAsnSerThrGluIleArgValAlaArgIleGlnGluIleCysGlyArgArgMetAsn 209
    |||
Db 1258 TCCAAATTTCTACTAAATACCGCTGACCGCAATTCAGAAATATGTGACGACGACAAATGAAT 1317
QY 210 ValSerIleThrLysLysAspPheProGluSerLeuProaspAlaaspAspArgGln 229
    |||
Db 1318 GTATCTATCAACAAGAAAGACTTCCGAAATCGCTTCCAGATGACATGACGACAGCA 1377
QY 230 ValPheSerSerGlySerGlnTyrTrpIleAspGluGlySerPheIleIleProVal 249
    |||
Db 1378 GTCTTTAGGCTCTGATCTCAATATGTGATGAGAGAGGAAGCTTTCATATTTCTGTT 1437
QY 250 GlyPheAlaAlaValAsnGlyTyrGlnLeuAsnAlaLysLysGlnTyrIleGluHisThr 269
    |||
Db 1438 GGAATTGACAGCATGCAATGATATCAACTAAATGCCAAAAGAAATATATTGACACACA 1497
QY 270 AsnLysIleAlaGlnAlaIleLysAsnGlyGluAsnProArgTyrAspSerAspAspVal 289
    |||
Db 1498 AATTAATTTGCTCAAGCAATTAATAAATGAGAAATCCAAATGATATGCTCAGACGATC 1557
QY 290 ThrPheAspGlnLeuAlaLysAspProIleAspPromeIleTyrParGlyValLysVal 309
    |||
Db 1558 ACATTTGATCAATTAACAAAAGATCCAAATTTGATGATGATTTGGAGAAAAGTTAAGGTT 1617
QY 310 GlyGlnLysPheGlnLeuIleAspProLeuAlaGlnInPheAsnAsnLeuHisValAla 329
    |||
Db 1618 GGCACAAAAGTTTACGCTCATGCAACCCCTTGCTCAGCAATTCATTAACCTCCACGTCGCT 1677
QY 330 SerIleLeuLysPheCysLysThrGlnGlyTyrLeuIleValGlyMetAspGlyProasp 349
    |||
Db 1678 TCGATTTCTCAAAATTTGCAAAAGTGAAGATATCTTATTTGGGAATGAGATGTCAGAT 1737
QY 350 AlaLeuGlnLysPhePheProIleHisIleAsnAsnThrPheMetPheProValGlyTyr 369
    |||
Db 1738 GCACCTTGAAAGAGCTTTTCTATTCATATCAATAAATACATTTATGTTCCAGTTGTTAT 1797
QY 370 AlaGlyLysTyrAsnLeuGlnLeuValAlaPropAspGluPheLysGlyThrPheArgTrp 389
    |||
Db 1798 GCGGAAAAGATATTAATTTGGAACTTGTCCGCCAGATGATGATTCAAAGAACATTCAGATGG 1857
QY 390 AspGluTyrLeuGlnLysGlnSerAlaGluThrLeuProLeuAspLeuPheLysPromeI 409

```

Db 1858 GATGAATACCTGGAGAAAGAAATCGCAGAAACCCGCTGACCTTTCAGGCCAATG 1917
Oy 410 Prosergingluatgleuaspserpshelysvalileuileserlysarval 427
Db 1918 CCTTCCCAAGAGAGATTAGACAAATTTAGGTAATTCGATTTCGAAACGGGTT 1971

RESULT 6
US-09-872-523-77
; Sequence 77, Application US/09872523
; Patent No. US20020137906A1
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Davison, Ewa M.
; APPLICANT: Lu, Xiaowei
; TITLE OF INVENTION: A Tumor Suppressor Pathway In C. Elegans
; FILE REFERENCE: 01997/536002
; CURRENT APPLICATION NUMBER: US/09/872,523
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/208,802
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 77
; LENGTH: 2307
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-872-523-77

Alignment Scores:
Pred. No.: 2,84e-287 Length: 2307
Score: 278.00 Matches: 278
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 55.82% Indels: 0
DB: 10 Gaps: 0

US-09-872-523-5 (1-498) x US-09-872-523-77 (1-2307)

Oy 150 Asncysileaspjglyluilevaliglylnhrserleuserprolyspheaspjgly 169
Db 1138 AATTCGATTCGATGGGAAATGTCGGCCAACTTCGCTGCCAAATTCGATCAAGG 1197
Oy 170 LysalaleuLeuserlyshisargpshelysvaliglylnatrgleuLeuLeuasnry 189
Db 1198 AAGGCTCTCCTAAGCAACATCGTTTCAAGTTGGACAACGCTTCAATTAATATAT 1257
Oy 190 Serasnserrhngluilearvalalargtiegngluilecysglyarargmetasn 209
Db 1258 TCCAATTCCTGAAATAGCGGTAGCGGCAATTCAGAAATATGTGGACGACGATGANT 1317
Oy 210 ValserlierhlyllyspPheProgluserleuProaspPalaaspPaspargln 229
Db 1318 GTATCTATCACAAGAAAGAACTTCCGAAATCGCTTCAGATCGAGATGACGACACAA 1377
Oy 220 ValPheSerSerGlySerGlnTyrTrpIleaspjglyserPhePheProval 249
Db 1378 GTCTTACCTCTGATCTCAATATTTGATAGACGAGGAGAGCTTCTCATATTTCTCTGT 1437
Oy 250 GlyPheAlaIalvalAsnGlyTyrGlnLeuasnAlalysGlnTyrIleGlnIshtr 269
Db 1438 GGATTTGACGACGATCAATGATATCACTAAATGCGAAAGAAAGAAATATTGACGACACA 1497
Oy 270 AsnlylIeAlaIalIleLysasnGlyLusnProArgTyrAspSerAspaspval 289
Db 1498 AATAAAATTCCTCAAGCAATATAAAAAATGAGAAAAATCCAGATATGATCAGACACACTC 1557
Oy 290 ThrPheaspGlnLeuAlaLysaspProIleaspProMetIleTPrargLysVallyVal 309
Db 1558 ACATTGTGATCAATTGCAAAAGATCCCAATGATCCCATGATTTGGAGAAAAGTTAAGGTT 1617
Oy 310 GlyGlnLysPheGlnLeuIleaspProLeuAlaGlnGlnPheAsnAsnLeuHisValAla 329

Db 1618 GGACAAAAGTTGAGCTCATGACCCCTTGCTCAGCAATTCATACCTCCACGCTGCT 1677
Oy 330 SerileuLysPheCysLysThrGluGlyTyrIleuilevaliglymetaspjlyProasp 349
Db 1678 TCGATTCCTCAATTTTGGAAAACTGAGAGATATCTTATTTGGGAATGATGATGCCAAT 1737
Oy 350 AlaLeuGluaspSerPheProIleHisIleasnshnrPhemetPheProvalGlyTyr 369
Db 1738 GCACCTGAAGACAGATTTCCTATCATATCATATATATATATATATATATATATATAT 1797
Oy 370 AlaGlyLysTyrAsnLeuGlnLeuValProProaspjlyPheLysGlyThrPheArgTrp 389
Db 1798 GCGGAAAAGTATATTTGAACTTGTTCGCCAGATGAGTTCAAGGAAATTCAGATGG 1857
Oy 390 AspGlyTyrLeuGluLysGlnSerAlaGluThrLeuProLeuAspLeuPheLysProMet 409
Db 1858 GATGAATACCTGGAGAAAGAAATTCGACAAACCCGACGCTTGACTTTCAGGCCAATG 1917
Oy 410 Prosergingluatgleuaspserpshelysvalileuileserlysarval 427
Db 1918 CCTTCCCAAGAGAGATTAGACAAATTTAGGTAATTCGATTTCGAAACGGGTT 1971

RESULT 7
US-09-872-523-78
; Sequence 78, Application US/09872523
; Patent No. US20020137906A1
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Davison, Ewa M.
; APPLICANT: Lu, Xiaowei
; TITLE OF INVENTION: A Tumor Suppressor Pathway In C. Elegans
; FILE REFERENCE: 01997/536002
; CURRENT APPLICATION NUMBER: US/09/872,523
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/208,802
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 78
; LENGTH: 2307
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-872-523-78

Alignment Scores:
Pred. No.: 2,84e-287 Length: 2307
Score: 278.00 Matches: 278
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 55.82% Indels: 0
DB: 10 Gaps: 0

US-09-872-523-5 (1-498) x US-09-872-523-78 (1-2307)

Oy 150 Asncysileaspjglyluilevaliglylnhrserleuserprolyspheaspjgly 169
Db 1138 AATTCGATTCGATGGGAAATGTCGGCCAACTTCGCTGCCAAATTCGATCAAGG 1197
Oy 170 LysalaleuLeuserlyshisargpshelysvaliglylnatrgleuLeuLeuasnry 189
Db 1198 AAGGCTCTCCTAAGCAACATCGTTTCAAGTTGGACAACGCTTGAATCATTAATATAT 1257
Oy 210 ValserlierhlyllyspPheProgluserleuProaspPalaaspPaspargln 229
Db 1258 TCCAATTCCTGAAATAGCGGTAGCGGCAATTCAGAAATATGTGGACGACGACAAAT 1317
Oy 220 ValserlierhlyllyspPheProgluserleuProaspPalaaspPaspargln 229
Db 1318 GTATCTATCACAAGAAAGAACTTCCGAAATCGCTTCAGATCGAGATGACGACACAA 1377
Oy 230 ValPheSerSerGlySerGlnTyrTrpIleaspjglyserPhePheProval 249
Db 1378 GTCTTACCTCTGATCTCAATATTTGATAGACGAGGAGACCTTCTCATATTTCTCTGT 1437


```
; SEQ ID NO 226880
; LENGTH: 590
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-226880

Alignment Scores:
Pred. No.:      67.7      Length:      590
Score:          8.00      Matches:      8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    1.61%      Indels:      0
DB:             13      Gaps:          0

US-09-872-523-5 (1-498) x US-10-027-632-226880 (1-590)

OY      159 GlnHrSerLeuSerProLysPhe 166
        |||||||
DB      157 CAACGTCCTTTCACCTAAATTC 180

RESULT 11
US-10-027-632-226881
; Sequence 226881, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: fastSeq for Windows Version 4.0
; SEQ ID NO 226881
; LENGTH: 590
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-226881

Alignment Scores:
Pred. No.:      67.7      Length:      590
Score:          8.00      Matches:      8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    1.61%      Indels:      0
DB:             13      Gaps:          0

US-09-872-523-5 (1-498) x US-10-027-632-226881 (1-590)

OY      159 GlnHrSerLeuSerProLysPhe 166
        |||||||
DB      157 CAACGTCCTTTCACCTAAATTC 180

RESULT 12
US-10-027-632-222125
; Sequence 222125, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: fastSeq for Windows Version 4.0
; SEQ ID NO 222125
; LENGTH: 608
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-222125

Alignment Scores:
Pred. No.:      69.6      Length:      608
Score:          8.00      Matches:      8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    1.61%      Indels:      0
DB:             13      Gaps:          0

US-09-872-523-5 (1-498) x US-10-027-632-222125 (1-608)

OY      6 LysileValArgAlaasnLysLys 13
        |||||||
DB      66 AAAATTGTTAGAGCTAATAAAAA 89

RESULT 13
US-10-027-632-6153
; Sequence 6153, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: fastSeq for Windows Version 4.0
; SEQ ID NO 6153
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-6153

Alignment Scores:
Pred. No.:      69.7      Length:      609
```

```
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: fastSeq for Windows Version 4.0
; SEQ ID NO 222125
; LENGTH: 608
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-222125

Alignment Scores:
Pred. No.:      69.6      Length:      608
Score:          8.00      Matches:      8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    1.61%      Indels:      0
DB:             13      Gaps:          0

US-09-872-523-5 (1-498) x US-10-027-632-222125 (1-608)

OY      6 LysileValArgAlaasnLysLys 13
        |||||||
DB      66 AAAATTGTTAGAGCTAATAAAAA 89

RESULT 13
US-10-027-632-6153
; Sequence 6153, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: fastSeq for Windows Version 4.0
; SEQ ID NO 6153
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-6153

Alignment Scores:
Pred. No.:      69.7      Length:      609
```



```
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.61% Indels: 0
DB: 13 Gaps: 0

US-09-872-523-5 (1-498) x US-10-027-632-6153 (1-609)
QY 213 ThrlyslvAspphProcluser 220
DB 377 ACCAAGAGGACTCCGAGTCT 400

RESULT 14
US-10-027-632-224608
; Sequence 224608, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; POLYMORPHISMS IN THE HUMAN GENOME
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 224608
; LENGTH: 611
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-224608

Alignment Scores:
Pred. No.: 70 Length: 611
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.61% Indels: 0
DB: 13 Gaps: 0

US-09-872-523-5 (1-498) x US-10-027-632-224608 (1-611)
QY 177 ArgPhelysValGlyGlnArgLeu 184
DB 159 AGGTCGAAGTGCGGCGAGACTG 182

RESULT 15
US-10-027-632-224609
; Sequence 224609, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; POLYMORPHISMS IN THE HUMAN GENOME
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
```

```
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 224609
; LENGTH: 611
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-224609

Alignment Scores:
Pred. No.: 70 Length: 611
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.61% Indels: 0
DB: 13 Gaps: 0

US-09-872-523-5 (1-498) x US-10-027-632-224609 (1-611)
QY 177 ArgPhelysValGlyGlnArgLeu 184
DB 159 AGGTCGAAGTGCGGCGAGACTG 182
```

Search completed: September 2, 2003, 23:19:00
Job time : 400 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 2, 2003, 19:33:40 ; Search time 5500 Seconds
(without alignments)
3704.181 Million cell updates/sec

Title: US-09-872-523-5

Perfect score: 498
Sequence: 1 MSEFLKIVRANKSRDKLKD.....PIGWCEAHNSYVLQPPKKNY 498

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 segs, 20454813386 residues

Word size: 1

Total number of hits satisfying chosen parameters: 5773148

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=x1h
-O/cqnt2.1/USPTO.spool/US09872523/runat.02092003.073001.27220/app-query.fasta.1.647
-DB=GenEmbl -OPMT=fastap -SUFFIX=oli.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptio
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09872523.@CEN.1.1.3508-@runat.02092003.073001.27220 -NCPU=6 -ICPU=3
-NO_MMP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database:

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_htg_hum:*
31: em_htg_in:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pin:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	278	55.8	30911	3 CER06C7	271266 Caenorhabd
C 2	11	2.2	99298	AC0090416	AC0090416 Caenorhab
C 3	10	2.0	142278	AC005177	AC005177 Homo sapi
C 4	10	2.0	182051	AC007638	AC007638 Homo sapi
C 5	9	1.8	74224	AC021724	AC021724 Homo sapi
C 6	9	1.8	98569	AC110778	AC110778 Homo sapi
C 7	9	1.8	110000	AC123272_0	AC123272 Rattus no
C 8	9	1.8	115532	AC008424	AC008424 Homo sapi
C 9	9	1.8	133322	AP005306	AP005306 Oryza sat
C 10	9	1.8	140542	AC079917	AC079917 Homo sapi
C 11	9	1.8	142623	AP005291	AP005291 Oryza sat
C 12	9	1.8	149425	AC013320	AC013320 Homo sapi
C 13	9	1.8	154229	AC008856	AC008856 Homo sapi
C 14	9	1.8	157102	AC016409	AC016409 Homo sapi
C 15	9	1.8	162649	BX332824	BX332824 Dario rer
C 16	9	1.8	166447	AC021723	AC021723 Homo sapi
C 17	9	1.8	169624	AC024187	AC024187 Homo sapi
C 18	9	1.8	170882	AC107839	AC107839 Mus muscu
C 19	9	1.8	171777	AC084713	AC084713 Homo sapi
C 20	9	1.8	172442	AC117530	AC117530 Homo sapi
C 21	9	1.8	177807	AC010411	AC010411 Homo sapi
C 22	9	1.8	181435	AC024244	AC024244 Homo sapi
C 23	9	1.8	191754	AC021016	AC021016 Homo sapi
C 24	9	1.8	233711	AC098070	AC098070 Rattus no
C 25	9	1.8	232053	AC097148	AC097148 Rattus no
C 26	8	1.6	330	AF103485	AF103485 Homo sapi
C 27	8	1.6	461	G75687	G75687 99-62654 Pe
C 28	8	1.6	576	11 HSPH25G5	AL159056 H.sapiens
C 29	8	1.6	657	11 BV075348	BV075348 S21P6643
C 30	8	1.6	795	11 BV036204	BV036204 S21P6012
C 31	8	1.6	826	11 BV067302	BV067302 S21P6686
C 32	8	1.6	872	6 BD146582	BD146582 primer to
C 33	8	1.6	974	3 AF099109	AF099109 Leishman
C 34	8	1.6	1442	6 BD073001	BD073001 70 human
C 35	8	1.6	1497	8 AB072454	AB072454 Coprinops
C 36	8	1.6	1830	9 HSM800595	AL080093 Homo sapi
C 37	8	1.6	2202	6 BD159393	BD159393 primer to
C 38	8	1.6	2202	9 AK022163	AK022163 Homo sapi
C 39	8	1.6	2271	9 HUWTLTIT	M99436 Human trans
C 40	8	1.6	2588	9 BC017364	BC017364 Homo sapi
C 41	8	1.6	2719	6 BD155878	BD155878 primer to
C 42	8	1.6	2719	9 AK001086	AK001086 Homo sapi
C 43	8	1.6	2738	9 AK009506	AK009506 Homo sapi
C 44	8	1.6	3524	6 AX409557	AX409557 Sequence
C 45	8	1.6	3524	6 BD080538	BD080538 Chemokine

RESULT 1

ALIGNMENTS

LOCUS	30911 bp	DNA	linear	INV 21-MAY-2003
DEFINITION	Caenorhabditis elegans cosmid R06C7, complete sequence.			
ACCESSION	J71266			
VERSION	J71266.1 GI:1279324			
KEYWORDS	HMG; Adenylsuccinate lyase; Myosin heavy chain; Rat T0AD-64 protein like; Serine/threonine-Protein kinase; Zinc finger protein. Caenorhabditis elegans			
SOURCE	Caenorhabditis elegans			
ORGANISM	Eukaryote; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.			
REFERENCE	1 none.			
AUTHORS	Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium			
TITLE	Science 282 (5396), 2012-2018 (1998)			
JOURNAL	99069613			
MEDLINE	9851916			
PUBMED	The C.elegans Sequencing Consortium.			
REMARK	2 (bases 1 to 30911)			
REFERENCE	Gardner,A.E.			
AUTHORS	Direct Submission			
TITLE	Submitted (19-APR-1996) Nematode Sequencing Project, Sanger			
JOURNAL	Institute, Hinxton, Cambridge CB10 1SA, England and Department of			
	Genetics, Washington University, St. Louis, MO 63110, USA. E-mail:			
	jesse@sanger.ac.uk or rwhematode.wustl.edu			
	Coding sequences below are predicted from computer analysis, using			
	predictions from GeneFinder (P. Green, U. Washington), and other			
	available information.			
	Current sequence finishing criteria for the C. elegans genome			
	sequencing consortium are that all bases are either sequenced			
	unambiguously on both strands, or on a single strand with both a			
	dye primer and dye terminator reaction, from distinct subclones.			
	Exceptions are indicated by an explicit note.			
	IMPORTANT: This sequence is not the entire insert of clone R06C7.			
	It may be shorter because we only sequence overlapping sections			
	once, or longer because we arrange for a small overlap between			
	neighbouring submissions.			
	The true left end of clone R06C7 is at 1 in this sequence. The true			
	right end of clone R06C7 is at 8718 in			
	sequence J71261.			
	The true left end of clone F21C3 is at 30808 in this sequence. The			
	start of this sequence (1..106) overlaps with the end of sequence			
	J298261.			
	The end of this sequence (30808..30911) overlaps with the start of			
	sequence J71261.			
	For a graphical representation of this sequence and its analysis			
	see: - http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?			
	name=R06C7			
	IMPORTANT: This sequence is NOT necessarily the entire insert of			
	the specified clone. It may be shorter because we only sequence			
	overlapping sections once, or longer because we arrange for a small			
	overlap between neighbouring submissions.			
	Location/Qualifiers			
	1..30911			
	/organism="Caenorhabditis elegans"			
	/mol.type="genomic DNA"			
	/strain="Bristol N2"			
	/db_xref="taxon:6239"			
	/chromosome="I"			
	/clone="R06C7"			
	join(complement(588..775),complement(395..539),			
	complement(169..295),complement(298261.1:204..368),			
	complement(271258.1:40294..40359),			
	complement(271258.1:40101..40244),			
	complement(271258.1:39857..39944),			
	complement(271258.1:39258..39701),			
	complement(271258.1:38262..39212),			
	complement(271258.1:38105..38207),			
	complement(271258.1:37940..38057),			

cds
 join(complement(588..775),complement(395..539),
 complement(169..295),complement(498261.1:204..366),
 complement(298261.1:105..155),
 complement(271258.1:40294..40359),
 complement(271258.1:40101..40244),
 complement(271258.1:39857..39944),
 complement(271258.1:39258..39701),
 complement(271258.1:38262..39212),
 complement(271258.1:38105..38207),
 complement(271258.1:37940..38057),
 complement(271258.1:37773..37855))
 /gene="C01H6.9"
 /standard_name="C01H6.9"
 /note="contains similarity to Pfam domain: PF00066
 (Eukaryotic protein kinase domain), Score=37.5,
 E-value=8.2e-10, N=2
 cDNA EST yk42h5.3 comes from this gene
 cDNA EST yk42h5.5 comes from this gene"
 /codon_start=1
 /product="Hypotheetical protein C01H6.9"
 /protein_id="CAA95847.1"
 /db_xref="GI:3878842"
 /db_xref="GOA:O17985"
 /db_xref="SPTRMBL:O17985"
 /translation="MPPKRIKIVIPISPKKANKNFARRNRQLSSVITTECPIDPEYE
 ESKNALFESFNDPEPEPKSKNNKRPYEPSPNVDSIESDEEDGQFVPVADA
 DEEETKIKKEAFSDLDVSLIEVFNNDIGRVINENKGIHKOGEARPEPKID
 PAKKMEORAKRVKQYPEPKNSANOVOCPTNASCSCITLADHMOENKASIALSFL
 RVFVFPQRKAKVPYPAKHVOCQDERLAITSYTPIRAKTKAQDIDTIGSYSPKQNA
 SPIPEKQVGRRRKSPFVVDVSLNDSCEINTIRKEQIALKNSAQKEETTINITPE
 ESKVRHESLAGSIQPOODLEETIASNRLSHSLKTPYIESRNESHQNSKIMTQ
 VSDSEVEVDNTNRAKETSVTLEHDEHERCSLDLEESYSKRRQKTSPEMQC
 GASTVMEVDQDDELALFEIKENETIKRTTLPOORPSRRHSINSAMEMSIQ
 FLIDTRGEERKQYSESRAESRNTPTGWTINEDSLPYLEDITSIDPSMMOL
 LHYVGESEKTSVDLSKASLDGRVNAKKEGAYGEYFTIMGKPAIKITVPEPDC
 NRPYEEYISEEMQTSVDVLPVETVAKLESLADDEAWSTNFTMTSAEVMGKYP
 KGLISAMDYSIDLESEENTRPDYSSIDONFTLPVANGIALIEDVLESENELSII
 HGVLSMADEALEFEHNDLHGNVLIDRNCVKELEYVHQKVPILSYHGKIVNID
 FTLSRISKATTVYMDLENDPAIFEGODDPOFEVYEMKRNCKSNMKSRRRTNMI
 VYANLIDITKICPKRLITKRRMEKLVFLDRAEFGSGCESLTINEPFSDEYEGPIG
 MSTRRO"
 complement(join(1315..1480,1528..1715,1747..1783,
 1834..2156))
 /gene="R06C7.2"
 complement(join(1315..1480,1528..1715,1747..1783,
 1834..2156))
 /gene="R06C7.2"
 /standard_name="R06C7.2"
 /product="Hypotheetical protein R06C7.2"
 /protein_id="CAA95841.1"
 /db_xref="GI:3878836"
 /db_xref="SPTRMBL:O21772"
 /translation="MPSAVPDRRCDCKRYRSQTSNONYWLKISEPNLGGKRIKG
 NFHLHAERKVLFGVSMLKDQPAKTEKROETTCOLSIANEENNRKLDOETEN
 KLRNLINSELKTVSSPENPLQIILFELISEDELEVEFHCAETPKLSPQLLF
 EPPVAVNAYKASGRVNNMEATILDAIPEVSGPEAARAKQITDSIKYESGFLIDR
 VNASALIIQSCSINSENF"
 join(3063..3227,3274..3854,3899..4631,4680..4855,
 4899..5074,5117..5463,5510..5749,5790..6097,6143..6254)
 /gene="R06C7.1"
 join(3063..3227,3274..3854,3899..4631,4680..4855,
 4899..5074,5117..5463,5510..5749,5790..6097,6143..6254)
 /gene="R06C7.1"
 /standard_name="R06C7.1"
 /note="contains similarity to Pfam domain: PF02170 (ZAP
 domain), Score=151.7, E-value=4.2e-42, N=1, PF02171 (P1w1
 domain), Score=456.7, E-value=6.3e-134, N=1
 cDNA EST yk31a12.5 comes from this gene
 cDNA EST yk31a12.3 comes from this gene
 cDNA EST yk31a12.3 comes from this gene
 cDNA EST yk31a12.3 comes from this gene
 cDNA EST yk36g4.3 comes from this gene

e-mail: jspeith@watson.wustl.edu

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

FEATURES

source

Location/Qualifiers

1..99298

/organism="Caenorhabditis briggsae"

/mol_type="genomic DNA"

/strain="Gujarat G16"

/db_xref="taxon:6238"

/clone="CB042007"

/complement(67280..67362)

/product="tRNA-Lys"

/note="codon recognized: AAG"

tRNA

BASE COUNT

29971 a 19350 c 19512 g 30465 t

ORIGIN

Alignment Scores:

Pred. No.:

1.77 Length: 99298

Score:

11.00 Matches: 11

Percent Similarity:

100.00% Conservative: 0

Best Local Similarity:

100.00% Mismatches: 0

Query Match:

2.21% Indels: 0

DB:

3 Gaps: 0

US-09-872-523-5 (1-498) x AC090416 (1-99298)

Oy 430 ArgfengluAlaAspmCysGluasnGln 440

Db 17516 CGTCTCGAAGCAGCATATGTGCGAGAACCA 17484

RESULT 3

AC005177/c

LOCUS Homo sapiens chromosome 17, clone Hc1T462L7, complete sequence.

DEFINITION

AC005177

AC005177.1

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

FitzHugh, W., Forrest, C., Funke, R., Gage, D., Gardyna, S., Gensheimer, S., Geraghty, K., Gilmartin, T., Grant, G., Hagos, B., Harris, K., Horton, L., Howland, J. C., Hui, L., Jacotot, L., Kann, L., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., Meldrim, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J., Nachman, A., Nahf, R., Naylor, J., O'Connor, T., Pavlin, B., Peterson, K., Riley, R., Roberts, D., Rossello, R., Roy, A., Shyam, R., Strange-Thomann, N., Stillwell, J., Stojanovic, N., Stone, C., Strickland, C., Subramanian, A., Torruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A., Wang, B., Wheeler, J., Wu, Y., Ye, W. J., Zhao, J. and Zody, M.

Direct Submission

Submitted (02-JUL-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 142278)

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

Repeat-Region

```

repeat_region      /rpt_family="MSTa"
                    11395..11462
repeat_region      /rpt_family="L2"
                    complement(11481..11786)
repeat_region      /rpt_family="Aluuo"
                    11897..12011
repeat_region      /rpt_family="L2"
                    complement(12056..12230)
repeat_region      /rpt_family="L2"
                    complement(12527..12575)
repeat_region      /rpt_family="L2"
                    12736..12784
repeat_region      /rpt_family="L1MA9"
                    12785..12829
repeat_region      /rpt_family="(CA)n"
                    12832..12939
repeat_region      /rpt_family="L1MA9"
                    complement(13173..13236)
repeat_region      /rpt_family="AT_rich"
                    complement(13356..13411)
repeat_region      /rpt_family="(CANA)n"
                    13432..13563
repeat_region      /rpt_family="MIR"
                    1605..16651
repeat_region      /rpt_family="AT_rich"
                    complement(17420..17487)
repeat_region      /rpt_family="L2"
                    complement(17594..17783)
repeat_region      /rpt_family="L2"
                    complement(17845..17927)
repeat_region      /rpt_family="L2"
                    complement(17982..18358)
repeat_region      /rpt_family="L2"
                    18409..18442
repeat_region      /rpt_family="AT_rich"
                    complement(18507..18531)
repeat_region      /rpt_family="AT_rich"
                    complement(18534..18656)
repeat_region      /rpt_family="L1PA5"
                    complement(19200..19268)
repeat_region      /rpt_family="(TA)n"
                    complement(22343..22363)
repeat_region      /rpt_family="AT_rich"
                    23186..23485
repeat_region      /rpt_family="AluSc"
                    complement(23963..24001)
repeat_region      /rpt_family="AT_rich"
                    complement(24260..24657)
repeat_region      /rpt_family="MLT1B"
                    24962..25196
repeat_region      /rpt_family="L1ME"
                    25382..25999
repeat_region      /rpt_family="L1ME"
                    26004..26194
repeat_region      /rpt_family="L1ME2"
                    complement(26565..27016)
repeat_region      /rpt_family="L1ME1"
                    complement(27250..27297)
repeat_region      /rpt_family="AT_rich"
                    complement(27728..27780)
repeat_region      /rpt_family="L2"
                    complement(28184..28481)
repeat_region      /rpt_family="AluSc"
                    complement(28496..28793)
repeat_region      /rpt_family="AluSc"
                    complement(28802..28901)
repeat_region      /rpt_family="(GA)n"
                    29017..29145
repeat_region      /rpt_family="FLAM_A"
                    29164..29190
repeat_region      /rpt_family="(GGGA)n"
                    29394..29482
repeat_region      /rpt_family="MLT2E"

```

```

repeat_region      29432..29649
                    /rpt_family="MLT1D"
repeat_region      29631..29808
                    /rpt_family="MLT1E"
repeat_region      complement(29957..30889)
                    /rpt_family="L1MA7"
repeat_region      complement(31345..31587)
                    /rpt_family="Aluub"
repeat_region      complement(31759..32055)
                    /rpt_family="AluSc"
repeat_region      complement(32171..32295)
                    /rpt_family="L2"
repeat_region      complement(32327..33832)
                    /rpt_family="L2"
repeat_region      complement(35067..35113)
                    /rpt_family="AT_rich"
repeat_region      35499..35737

```

```

Alignment Scores:
Pred. No.:          34.6          Length:      142278
Score:             10.00          Matches:      10
Percent Similarity: 100.00%       Conservative: 0
Best Local Similarity: 100.00%    Mismatches:  0
Query Match:       2.01%          Indels:      0
DB:                9              Gaps:         0

```

US-09-872-523-5 (1-498) x AC005177 (1-142278)

QY 243 Serphei1epheprovalglyphea1a 252

Db 130587 TCTTTTTCATCTTCCTGCTGGTTCGCC 130558

```

RESULT 4
AC007638      182051 bp      DNA      linear      PRI 02-DEC-2001
LOCUS         Homo sapiens chromosome 17, clone RP11-515017, complete sequence.
DEFINITION    AC007638
ACCESSION     AC007638
VERSION       AC007638.8  GI:17226706
KEYWORDS      HTG.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 182051)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 17, clone RP11-515017
Unpublished
2 (bases 1 to 182051)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Depayre,E., Devon,K., Dewar,K.,
Donelan,L., Doyle,M., Ferreira,P., Fitzhugh,W., Forrest,C.,
Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
Karatas,A., Lehoczyk,J., Lieu,C., Locke,K., Macdonald,P.,
Mardquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
Peterson,K., Pollara,V., Riley,R., Roberts,S., Roy,A., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
Tesfaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
Direct Submission
Submitted (22-MAY-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 182051)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhalter,B.,
Brown,A., Camarata,J., Campotiano,A., Chang,J., Chazaro,B.,
Choepe1,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,

```

TITLE JOURNAL
REFERENCE
AUTHORS

REFERENCE
AUTHORS

Glnde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lmazares, R., Landers, T., Lenoczky, J., Levine, R., Liu, G., Maclean, C., MacDonald, P., Major, J., Marcusi, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (02-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Dec 2, 2001 this sequence version replaced gi:13958524.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L458
Center clone name: 515_O_17

FEATURES
source

Location/Qualifiers
1..182051
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-515017"
/clone_1lb="RPC1-11 Human Male BAC"
complement(1353..1468)
/rpt_family="L2"
complement(1805..2130)
/rpt_family="AluY"
2602..2760
/rpt_family="AluYb"
complement(3353..3656)
/rpt_family="AluX"
complement(4256..4412)
/rpt_family="L2"
4584..6080
/rpt_family="L1MEC"
6084..6439
/rpt_family="AluSg"
6502..6762
/rpt_family="AluX"
6855..7007
/rpt_family="L1MEC"
7008..7302
/rpt_family="AluSg"
7303..7592
/rpt_family="L1MEC"
7593..7697
/rpt_family="MIR"
7894..7983
/rpt_family="MIR3"
10083..10103
/rpt_family="AT_rich"
11175..11198
/rpt_family="AT_rich"
11941..12977
/rpt_family="L2"

repeat_region 13421..13462
/rpt_family="AT_rich"
repeat_region 13657..13961
/rpt_family="AluSg"
complement(14025..14134)
/rpt_family="L1MC/D"
complement(14156..14355)
/rpt_family="MIR"
15167..15343
/rpt_family="MIR"
complement(15613..15918)
/rpt_family="AluX"
15856..15885
/note="PCR reads only"
15886..16038
/note="PCR reads only"
complement(16039..16103)
/note="PCR reads only"
complement(16252..16336)
/rpt_family="L2"
complement(16879..17191)
/rpt_family="L2"
complement(17385..17797)
/rpt_family="MSTA"
complement(18004..18306)
/rpt_family="AluSg"
complement(18406..18850)
/rpt_family="L2"
18851..19257
/rpt_family="MER31B"
complement(19258..19308)
/rpt_family="L2"
complement(19543..20261)
/rpt_family="LTR34"
20409..20720
/rpt_family="AluY"
20940..21074
/rpt_family="MIR"
21287..21593
/rpt_family="L2"
complement(21594..21750)
/rpt_family="FRAM"
21751..22415
/rpt_family="L2"
22446..22805
/rpt_family="AluSg/X"
22815..23063
/rpt_family="L2"
complement(23803..23918)
/rpt_family="MIR"
complement(25013..25306)
/rpt_family="AluJo"
complement(25483..25688)
/rpt_family="L2"
25769..26010
/rpt_family="MLT1L"
complement(26304..26420)
/rpt_family="MIR"
26831..26851
/rpt_family="AT_rich"
complement(26852..27023)
/rpt_family="FRAM"
complement(27026..27326)
/rpt_family="AluJo"
complement(27337..27691)
/rpt_family="L1MC3"
27692..28054
/rpt_family="THE1B"
complement(28055..28551)
/rpt_family="L1MC3"
28601..28717
/rpt_family="MIR"
complement(29115..29419)


```

repeat_region /rpt_family="AluSg1"
                complement(29420..29553)
repeat_region /rpt_family="TH1B"
                complement(29554..29913)
repeat_region /rpt_family="TH1B"
                complement(29914..31486)
repeat_region /rpt_family="TH1B-int"
                complement(31487..31680)
repeat_region /rpt_family="TH1B"
                complement(31681..31976)
repeat_region /rpt_family="AluY"
                complement(31977..32154)
repeat_region /rpt_family="TH1B"
                complement(32155..32440)
repeat_region /rpt_family="TH1B"
                complement(32441..32736)
repeat_region /rpt_family="MIR"
                complement(32737..33840)

```

Alignment Scores:

```

Pred. No.: 43 Length: 182051
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.01% Indels: 0
DB: 9 Gaps: 0

```

US-09-872-523-5 (1-498) x AC007638 (1-182051)

Qy 243 SerPhePhePheProValGlyPheAla 252

Db 172894 TCTTTTTCATCTTCTCTGTGGGTTTCGCC 172923

```

RESULT 5
AC021724 AC021724 74224 bp DNA linear HTG 13-JUL-2000
LOCUS Homo sapiens chromosome 5 clone RP11-329C20 map 5, LOW-PASS
DEFINITION
SEQUENCE SAMPLING.

```

```

AC021724 AC021724.2 GI:9155264
VERSION HTG: HTGS_PHASE0.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
1 (bases 1 to 74224)
2 (bases 1 to 74224)
JOURNAL Homo sapiens chromosome 5, clone RP11-329C20
REFERENCE. Unpublished
AUTHORS

```

```

TITLE
JOURNAL
COMMENT
Submitted (19-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced g1:6715792.
All repeats were identified using RepeatMasker:
Smil, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research

```

```

Center code: WMBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L5961
Center clone name: 329_C_20

```

```

* NOTE: This record contains 83 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

```

```

1 783 782: contig of 782 bp in length
883 882: gap of 100 bp
1684 1684: contig of 802 bp in length
1784 1784: gap of 100 bp
1585 1585: contig of 800 bp in length
2585 2585: gap of 100 bp
3467 3467: contig of 783 bp in length
3567 3567: gap of 100 bp
4348 4348: contig of 781 bp in length
4448 4448: gap of 100 bp
5245 5245: contig of 797 bp in length
5345 5345: gap of 100 bp
5449 5449: gap of 100 bp
5246 5246: gap of 100 bp
5346 5346: contig of 755 bp in length
6100 6100: contig of 755 bp in length
6200 6200: gap of 100 bp
6301 6301: contig of 807 bp in length
7008 7008: gap of 100 bp
7107 7107: contig of 786 bp in length
7893 7893: gap of 100 bp
7894 7894: gap of 100 bp
7993 7993: contig of 812 bp in length
8805 8805: gap of 100 bp
8905 8905: gap of 100 bp
8917 8917: contig of 812 bp in length
9718 9718: gap of 100 bp
9817 9817: contig of 805 bp in length
10622 10622: gap of 100 bp
10722 10722: gap of 100 bp
11512 11512: contig of 790 bp in length
11613 11613: gap of 100 bp
12358 12358: contig of 746 bp in length
12458 12458: gap of 100 bp
13251 13251: contig of 793 bp in length
13351 13351: gap of 100 bp
14150 14150: contig of 799 bp in length
14250 14250: gap of 100 bp
15067 15067: contig of 817 bp in length
15167 15167: gap of 100 bp
15979 15979: contig of 812 bp in length
16079 16079: gap of 100 bp
16799 16799: gap of 100 bp
16847 16847: contig of 768 bp in length
16947 16947: gap of 100 bp
17666 17666: contig of 819 bp in length
17866 17866: gap of 100 bp
18681 18681: contig of 815 bp in length
18781 18781: gap of 100 bp
19591 19591: contig of 810 bp in length
19691 19691: gap of 100 bp
20478 20478: contig of 787 bp in length
20578 20578: gap of 100 bp
21356 21356: contig of 778 bp in length
21456 21456: gap of 100 bp
22219 22219: contig of 763 bp in length
22319 22319: gap of 100 bp
23108 23108: contig of 789 bp in length
23208 23208: gap of 100 bp
23981 23981: contig of 773 bp in length
24081 24081: gap of 100 bp
24886 24886: contig of 805 bp in length

```

```

* 24887 24986: gap of 100 bp
* 24987 25807: contig of 821 bp in length
* 25808 25907: gap of 100 bp
* 25908 26687: contig of 780 bp in length
* 26688 26787: gap of 100 bp
* 26788 27598: contig of 811 bp in length
* 27599 27698: gap of 100 bp
* 27699 28512: contig of 814 bp in length
* 28513 29425: contig of 813 bp in length
* 29426 29525: gap of 100 bp
* 29526 30301: contig of 776 bp in length
* 30302 30402: gap of 100 bp
* 30402 31185: contig of 784 bp in length
* 31186 32075: contig of 790 bp in length
* 32076 32175: gap of 100 bp
* 32176 32970: contig of 795 bp in length
* 32971 33070: gap of 100 bp
* 33071 33874: contig of 804 bp in length
* 33875 33974: gap of 100 bp
* 33975 34771: contig of 797 bp in length
* 34772 34871: gap of 100 bp
* 34872 35682: contig of 811 bp in length
* 35683 35782: gap of 100 bp
* 35783 36591: contig of 809 bp in length
* 36592 37463: contig of 772 bp in length
* 37464 37563: gap of 100 bp
* 37564 38365: contig of 802 bp in length
* 38366 38465: gap of 100 bp
* 38466 39264: contig of 799 bp in length
* 39265 39364: gap of 100 bp
* 39365 40166: contig of 802 bp in length
* 40167 40266: gap of 100 bp
* 40267 41046: contig of 780 bp in length
* 41047 41146: gap of 100 bp
* 41147 41945: contig of 799 bp in length
* 41946 42045: gap of 100 bp
* 42046 42841: contig of 796 bp in length
* 42842 42941: gap of 100 bp
* 42942 43751: contig of 810 bp in length
* 43752 43851: gap of 100 bp
* 43852 44656: contig of 805 bp in length
* 44657 44756: gap of 100 bp
* 44757 45564: contig of 808 bp in length
* 45565 45664: gap of 100 bp
* 45665 46480: contig of 816 bp in length
* 46481 46580: gap of 100 bp
* 46581 47397: contig of 817 bp in length
* 47398 47497: gap of 100 bp
* 47498 48300: contig of 803 bp in length
* 48301 48400: gap of 100 bp
* 48401 49208: contig of 808 bp in length
* 49209 49308: gap of 100 bp
* 49309 50099: contig of 791 bp in length
* 50100 50199: gap of 100 bp
* 50200 50961: contig of 764 bp in length
* 50962 51061: gap of 100 bp
* 51062 51869: contig of 808 bp in length
* 51870 51969: gap of 100 bp
* 51970 52775: contig of 806 bp in length
* 52776 52875: gap of 100 bp
* 52876 53675: contig of 800 bp in length
* 53676 53775: gap of 100 bp
* 53776 54577: contig of 802 bp in length
* 54578 54677: gap of 100 bp
* 54678 55451: contig of 774 bp in length
* 55452 55551: gap of 100 bp
* 55552 56348: contig of 797 bp in length
* 56349 56448: gap of 100 bp
* 56449 57260: contig of 812 bp in length
* 57261 57360: gap of 100 bp

```

```

* 57361 58165: contig of 805 bp in length
* 58166 58265: gap of 100 bp
* 58266 59057: contig of 792 bp in length
* 59058 59157: gap of 100 bp
* 59158 59920: contig of 763 bp in length
* 59921 60021: gap of 100 bp
* 60021 60802: contig of 782 bp in length
* 60803 60902: gap of 100 bp
* 60903 61632: contig of 729 bp in length
* 61632 62513: contig of 100 bp
* 61732 62513: contig of 782 bp in length
* 62513 62613: gap of 100 bp
* 62614 63418: contig of 805 bp in length

```

Alignment Scores:

```

Pred. No.: 277 Length: 74224
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.81% Indels: 0
DB: 2 Gaps: 0

```

US-09-872-523-5 (1-498) x AC021724 (1-74224)

```

Oy 359 Ileaasnsnrphmetpmetpova1 367
Db 70395 ATAAACACACCTTATGTTCCAGTC 70421

```

RESULT 6

AC110778/c

LOCUS Homo sapiens BAC clone RP11-320L7 from 4, complete sequence.

DEFINITION

AC110778

VERSION

AC110778..4 GI:19909444

KEYWORDS

HTG.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 98569)

Wang, C., and Kozlowicz, A.

The sequence of Homo sapiens BAC clone RP11-320L7

Unpublished (2001)

JOURNAL

3 (bases 1 to 98569)

Waterston, R.H.

Direct Submission

Submitted (15-FEB-2002) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

4 (bases 1 to 98569)

Waterston, R.H.

Direct Submission

Submitted (03-APR-2002) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

5 (bases 1 to 98569)

Waterston, R.H.

Direct Submission

Submitted (16-APR-2002) Department of Genetics, Washington

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

On Apr 3, 2002 this sequence version replaced gi:19551181.

COMMENT

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@atson.wustl.edu
----- Summary Statistics
Center project name: H.NH0320107

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPC1-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Taleno, M., Catanesi, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-55L3, 2000 bp overlap.
Actual start of this clone is at base position 87259 of RP11-55L3;
actual end is at base position 98569 of RP11-320L7.

Polymorphisms exist between AC110778 and AC106892.

Location/Qualifiers

1. 98569

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="4"

/map="4"

/clone="RP11-320L7"

/clone_lib="RPC1-11"

118. 414

/rpt_family="Alu"

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

12271. 12478

/rpt_family="Alu"

14082. 14132

/rpt_family="GA-rich"

14133. 14490

/rpt_family="(GGAA)n"

14491. 14684

/rpt_family="GA-rich"

14794. 15091

/rpt_family="Alu"

16245. 16373

/rpt_family="MIR"

16456. 16576

/rpt_family="L2"

18135. 18205

/rpt_family="(TTTG)n"

18375. 18836

/rpt_family="ERV1"

18957. 18991

/rpt_family="AT-rich"

19506. 19814

/rpt_family="Alu"

23246. 23345

/rpt_family="Alu"

24690. 24767

/rpt_family="AT-rich"

24839. 25133

/rpt_family="Alu"

25392. 25503

/rpt_family="MIR"

25898. 26823

/rpt_family="L1"

28099. 28199

/rpt_family="(TA)n"

28678. 28776

/rpt_family="MIR"

28867. 28908

/rpt_family="AT-rich"

28979. 29126

/rpt_family="MIR"

29127. 29437

/rpt_family="Alu"

29438. 29521

/rpt_family="MIR"

30197. 30218

/rpt_family="AT-rich"

33309. 33331

/rpt_family="AT-rich"

34266. 34310

/rpt_family="AT-rich"

34518. 34546

/rpt_family="AT-rich"

35972. 36227

/rpt_family="MIR"

39103. 39193

/rpt_family="AT-rich"

40775. 40825

/rpt_family="AT-rich"

41371. 41456

/rpt_family="MER2-type"

41518. 41614

/rpt_family="AT-rich"

41669. 41703

/rpt_family="AT-rich"

41885. 41928

/rpt_family="AT-rich"

42029. 42926

/rpt_family="L1"

43610. 43714

/rpt_family="BC200"

43715. 44087

/rpt_family="MIR"

44088. 44472

```

repeat_region      /rpt_family="MaLR"
                    44473..45591
                    /rpt_family="L1"
repeat_region      45786..46233
                    /rpt_family="MaLR"
repeat_region      46336..46374
                    /rpt_family="CRNA-1le-ATA"
repeat_region      46564..46621
                    /rpt_family="AT_rich"
repeat_region      46757..46788
                    /rpt_family="(TC)n"

Alignment Scores:
Pred. No.:          356
Score:              9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match:        1.81%
DB:                  9
Gaps:                0

US-09-872-523-5 (1-498) x AC110778 (1-98569)
Oy      186      LeuleuAentYrSerAserThrglu 194
DB      8364     TTGTTGAATTACTCTACTACTGAA 8338

RESULT 7
AC123272.0
MPCOMMENT
Sequence split into 4 fragments
Fragment Name      Begin      End
AC123272_0         1      110000
AC123272_1         100001   210000
AC123272_2         200001   310000
AC123272_3         300001   369520
LOCUS      AC123272      369520 bp      DNA      linear      HTG 08-OCT-2002
DEFINITION Rattus norvegicus clone CH230-308E15, *** SEQUENCING IN PROGRESS
            *** 10 unordered pieces.
ACCESSION   AC123272
VERSION     AC123272.3  GI:23269854
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
            1 (bases 1 to 369520)
            Munry,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Angiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalto,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,Y., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Detamo,C., Ding,Y., Dinh,H., Divya,K.,
            Diaper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregorjis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,M.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Huliy,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Kapachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorenshewa,L., Louiseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,

```

```

Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nnekelemeh,O., Okwunu,G., Olarnunsaqoon,A., Pal,S., Pairs,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Pioppert,F., Polidexter,A., Popovic,D., Plinius,E., Pu,L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajz,D.,
Sneed,A., Sodergren,E., Song,X.Z., Sorelle,R., Sosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingley,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,Y., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,O., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,Y.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 369520)
Worley,K.C.
Submitted (29-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 369520)
Rat Genome Sequencing Consortium.
Submitted (08-OCT-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Sep 23, 2002 this sequence version replaced gi:21671645.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GMVI
Center clone name: CH230-308E15
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 226252 bases at least Q40
Consensus quality: 230304 bases at least Q30
Consensus quality: 232703 bases at least Q20
Estimated insert size: 246893; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.

```

* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 48216: contig of 48216 bp in length
* 48217 48316: gap of unknown length
* 48317 302116: contig of 253800 bp in length
* 302217 302216: gap of unknown length
* 302217 316444: contig of 14228 bp in length
* 316445 316544: gap of unknown length
* 316545 317730: contig of 1186 bp in length
* 317731 317830: gap of unknown length
* 317831 319510: contig of 1680 bp in length
* 319511 321036: contig of 1426 bp in length
* 321037 321136: gap of unknown length
* 321137 324733: contig of 3597 bp in length
* 324734 324833: gap of unknown length
* 324834 327945: contig of 3112 bp in length
* 327946 328045: gap of unknown length
* 328046 343427: contig of 15382 bp in length
* 343428 343527: gap of unknown length
* 343528 369520: contig of 25993 bp in length.

FEATURES
source
1. .369520
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-308E15"
misc_feature
48317..49352
/note="wgs-contig"
misc_feature
231258..232287
/note="wgs-contig"
BASE COUNT 76044 a 44227 c 43334 g 70304 t 135611 others
ORIGIN

Alignment Scores:
Pred. No.: 392 Length: 110000
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.81% Indels: 0
DB: 2 Gaps: 0

US-09-872-523-5 (1-498) x AC123272_0 (1-110000)
Oy 58 GluGlyValIlePheGluThrValVal 66
Db 93002 GAGGGTGTCTTCATTGAGACTGTAGTT 93028
|||||

RESULT 8
AC008424 115532 bp DNA linear HTG 18-JUL-2000
LOCUS Homo sapiens chromosome 5 clone CTC-259F19, WORKING DRAFT SEQUENCE,
DEFINITION
6 ordered pieces.
AC008424
VERSION AC008424.5 GI:9255971
KEYWORDS HTG: HTGS_PHASE2; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 115532)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 5
Unpublished
2 (bases 1 to 115532)
DOE Joint Genome Institute.
Direct Submission
Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Jul 18, 2000 this sequence version replaced gi:7708847.
-----Genome Center
Center: Joint Genome Institute

Center Code: JGI
Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 300134
Center Clone Name: CIT-HSPC_299F19

Summary Statistics
Consensus quality: 112143 bases at least Q40
Consensus quality: 114467 bases at least Q30
Consensus quality: 114880 bases at least Q20
Estimated insert size: 119000; pulse field gel estimation
Estimated insert size: 115332; sum-of-contigs estimation
Quality coverage: 5.73 in Q20 bases; pulse field gel estimation
Quality coverage: 5.92 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

1 3722: contig of 3722 bp in length
* 3723 3822: gap of unknown length
* 3823 70532: contig of 66710 bp in length
* 70533 70632: gap of unknown length
* 70633 76281: contig of 5649 bp in length
* 76282 76381: gap of unknown length
* 76382 94828: contig of 18447 bp in length
* 94829 94928: gap of unknown length
* 94929 107602: contig of 12674 bp in length
* 107603 107702: gap of unknown length
* 107703 115532: contig of 7830 bp in length.

FEATURES
source
1. .115532
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTC-259F19"
/clone_lib="Caltech human BAC library C"
BASE COUNT 37053 a 20285 c 20614 g 37079 t 501 others
ORIGIN

Alignment Scores:
Pred. No.: 409 Length: 115532
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.81% Indels: 0
DB: 2 Gaps: 0

US-09-872-523-5 (1-498) x AC008424 (1-115532)
Oy 359 IleAsnThrPheMetPheProVal 367
Db 97280 ATAAACACACCTTATGTTCCAGTC 97306
|||||

RESULT 9
AP005306 133322 bp DNA linear HTG 31-MAY-2002
LOCUS Oryza sativa (japonica cultivar-group) chromosome 2 clone P0030D07,
DEFINITION
*** SEQUENCING IN PROGRESS ***
AP005306
VERSION AP005306.1 GI:21280356
KEYWORDS HTG: HTGS_PHASE2.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE
AUTHORS
TITLE
1
Sasaki, T., Matsumoto, T. and Katayose, Y.
Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, PAC
clone: P0030007

JOURNAL
REFERENCE
AUTHORS
TITLE
2
Published Only in Database (2002)
Sasaki, T., Matsumoto, T. and Katayose, Y.
Direct Submission
Submitted (29-MAY-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program, Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgrp.dna.affrc.go.jp/
Tel: 81-298-38-7441, Fax: 81-298-38-7468)

COMMENT
NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a working draft sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
Location/Qualifiers
1..133322
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="2"
/clone="P0030007"

BASE COUNT
38223 a 28234 c 28707 g 38158 t

ORIGIN

Alignment Scores:
Pred. No.: 465 Length: 133322
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.81% Indels: 0
DB: 2 Gaps: 0

US-09-872-523-5 (1-498) x AP005306 (1-133322)

OY
264 GUTYRTTGGUHSHTAASLSTLE 272
|||||

Db
26156 GAATATATAGAACATCAATAAATA 26130

RESULT 10
AC079917/c 140542 bp DNA linear PRI 09-APR-2002

LOCUS
AC079917 Homo sapiens chromosome 11, clone RP11-55122, complete sequence.

DEFINITION
AC079917
AC079917.6 GI:20087233
HTG.

VERSION
AC079917.6 GI:20087233

KEYWORDS
Homo sapiens (human)

SOURCE
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 140542)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 11, clone RP11-55122
Unpublished
2 (bases 1 to 140542)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Beda, F., Boguslavsky, L.,
Boukhgalter, B., Brown, A., Burkett, G., Campiano, A., Castle, A.,
Choe, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferreira, P.,
Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M.,
Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L.,
Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Larocque, K.,
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G.,

Macdonald, P., Marguis, N., McCarthy, M., McKernan, P., McKernan, K.,
McPheters, R., Meldrum, J., Meneus, L., Mihova, T., Mieng, V.,
Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T.,
O'Donnell, P., O'Neill, D., Olivari, T.M., Oliver, J., Peterson, K.,
Pierre, N., Pisan, C., Pollara, V., Raymond, C., Rieback, M., Riley, R.,
Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P.,
Sougnere, C., Spencer, B., Strange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Tirrell, A., Travers, M., Triggillo, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,
Zimmer, A. and Zody, M.

Direct Submission
Submitted (17-SEP-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 140542)

JOURNAL
REFERENCE
AUTHORS
1
Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhgalter, B., Brown, A., Camarata, J., Campiano, A., Chang, J.,
Chazaro, B., Choe, Y., Colangelo, M., Collins, S., Collymore, A.,
Cooke, P., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Larocque, K., Lamazares, R.,
Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Marguis, N., Matthews, C.,
McCarthy, M., McKernan, P., McKernan, K., Meldrum, J., Meneus, L.,
Mihova, T., Mieng, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Nobu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnuppock, R.,
Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (25-MAR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 140542)

JOURNAL
REFERENCE
AUTHORS
1
Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhgalter, B., Brown, A., Camarata, J., Campiano, A., Chang, J.,
Chazaro, B., Choe, Y., Colangelo, M., Collins, S., Collymore, A.,
Cooke, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Larocque, K., Lamazares, R.,
Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Marguis, N., Matthews, C.,
McCarthy, M., McKernan, P., McKernan, K., Meldrum, J., Meneus, L.,
Mihova, T., Mieng, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Nobu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnuppock, R.,
Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (09-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 9, 2002 this sequence version replaced gi:19703152.
ALL repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: LI0944
 Center clone name: 55_1_22

FEATURES

```

Source
-----
Location/Qualifiers
1..140542
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-55122"
/clone_lib="RPC1-11 Human Male BAC"
162..260
/rpt_family="L1MA4"
repeat_region
261..560
/rpt_family="AluY"
unsure
318..323
/note="<30 qual SINGL region"
561..770
/rpt_family="L1MA4"
repeat_region
890..929
/rpt_family="AT_rich"
repeat_region
1055..1157
/rpt_family="CT-rich"
repeat_region
complement(1627..2337)
/rpt_family="L1M4"
repeat_region
1980..1984
/note="<30 qual SINGL region"
2117..2122
/note="<30 qual SINGL region"
2153..2157
/note="<30 qual SINGL region"
2175..2180
/note="single clone coverage"
repeat_region
complement(2338..2630)
/rpt_family="AluSx"
repeat_region
complement(2631..4088)
/rpt_family="L1M4"
repeat_region
complement(4098..4475)
/rpt_family="L1M4"
repeat_region
4518..4627
/rpt_family="(TATG)n"
complement(4631..5487)
/rpt_family="L1ME1"
repeat_region
complement(5719..6081)
/rpt_family="L1MA7"
repeat_region
complement(6082..6390)
/rpt_family="AluJb"
repeat_region
complement(6391..6706)
/rpt_family="L1MA7"
repeat_region
complement(6709..7293)
/rpt_family="MSTB1-int"
repeat_region
complement(7298..7716)
/rpt_family="MSTB1"
repeat_region
complement(7717..8744)
/rpt_family="L1MA7"
repeat_region
complement(8745..9042)
/rpt_family="AluSx"
repeat_region
complement(9043..10284)
/rpt_family="L1MA7"
repeat_region
10336..10380
/rpt_family="L1MD2"
repeat_region
10392..10460
/rpt_family="(TTAA)n"
repeat_region
complement(10461..10858)
/rpt_family="L1ME1"
repeat_region
complement(10859..11227)
/rpt_family="L1ME1"
repeat_region
11234..11318
/rpt_family="(TTTA)n"

```

```

repeat_region complement(11350..12565)
/rpt_family="L1PA3"
repeat_region complement(12578..12979)
/rpt_family="L1ME1"
repeat_region complement(13067..13120)
/rpt_family="L1ME1"
repeat_region complement(13121..13429)
/rpt_family="AluY"
repeat_region complement(13430..14244)
/rpt_family="L1ME1"
repeat_region 14325..14360
/rpt_family="AT_rich"
repeat_region 14467..14485
/rpt_family="L1MD2"
repeat_region 14486..14792
/rpt_family="AluY"
repeat_region 14793..14919
/rpt_family="L1MD2"
repeat_region 14994..15372
/rpt_family="L1MD2"
repeat_region complement(15371..15704)
/rpt_family="L1M4"
repeat_region 15984..16425
/rpt_family="L1MD2"
repeat_region complement(16427..16477)
/rpt_family="MLT2B1"
repeat_region 16478..16503
/rpt_family="(CA)n"
repeat_region complement(16504..16662)
/rpt_family="MLT2B1"
repeat_region complement(16683..16883)

```

Alignment Scores:

Pred. No.:	487	Length:	140542
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.81%	Indels:	0
DB:	9	Gaps:	0

US-09-872-523-5 (1-498) x AC079917 (1-140542)

Qy 418 PheLysValIleLeuIleSerLysArg 426
 Db 41071 TTTAAGTTATATTAACTAAAGAA 41045

RESULT 11

AP005291/c

LOCUS

DEFINITION

AP005291

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AP005291 142623 bp DNA linear HTG 31-MAY-2002
 Oryza sativa (japonica cultivar-group) chromosome 2 clone
 OJ1282.H11, *** SEQUENCING IN PROGRESS ***.
 AP005291.1 GI:21280341
 HTG: HTGS_PHASE2.
 Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 Spermatophyta: Magnoliophyta: Liliopsida: Poales: Poaceae:
 Ehrhartoideae: Oryzaceae: Oryza.
 1
 Sasaki, T., Matsumoto, T. and Katayose, Y.
 Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, BAC
 clone: OJ1282.H11
 Published only in Database (2002)
 2 (bases 1 to 142623)
 Sasaki, T., Matsumoto, T. and Katayose, Y.
 Direct Submission
 Submitted (29-MAY-2002) Takuji Sasaki, National Institute of
 Agrobiological Sciences, Rice Genome Research Program, Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail: tsasaki@ias.affrc.go.jp, URL: <http://rgp.dna.affrc.go.jp/>,
 Tel: 81-298-38-7441, Fax: 81-298-38-7468)
 The nucleotide sequence of this BAC clone was generated by

combining Monsanto and RGP-Japan sequencing data.

NOTE: It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

* NOTE: This is a 'working draft' sequence.

* This sequence will be replaced

* by the finished sequence as soon as it is available and the accession number will be preserved.

FEATURES

Source

Location/Qualifiers
1. 142623
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="2"
/clone="OJ1282.H11"

BASE COUNT 40863 a 30430 c 31641 g 39639 t 50 others
ORIGIN

Alignment Scores:

Pred. No.:	493	Length:	142623
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.81%	Indels:	0
DB:	2	Gaps:	0

US-09-872-523-5 (1-498) x AP005291 (1-142623)

OY 264 GUTYTTTGGUHHSTHAsnlystle 272

DB 112946 GAATATATAGAACATCAATAAATA 112920

RESULT 12

AC013320/c

LOCUS AC013320 149425 bp DNA linear HTG 18-NOV-2000
DEFINITION Homo sapiens chromosome 2 clone RP11-36003 map 2, WORKING DRAFT
SEQUENCE, 19 unordered pieces.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

Project Information
Center project name: L3776
Center clone name: 360_O_3

Summary Statistics
Sequencing vector: M13; M77815; 12% of reads
Sequencing vector: Plasmid; n/a; 8% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 140887 bases at least Q40
Consensus quality: 144879 bases at least Q30
Consensus quality: 146417 bases at least Q20
Insert size: 228000; agarose-fp
Insert size: 147625; sum-of-contigs
Quality coverage: 3.6 in Q20 bases; agarose-fp
Quality coverage: 5.5 in Q20 ba.
NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 8144: contig of 8144 bp in length
* 8145 8244: gap of 100 bp
* 8245 9350: contig of 1106 bp in length
* 9351 9450: gap of 100 bp
* 9451 11335: contig of 1885 bp in length
* 11336 11435: gap of 100 bp
* 11436 14325: contig of 2890 bp in length
* 14326 14425: gap of 100 bp
* 14426 17305: contig of 2880 bp in length
* 17306 17405: gap of 100 bp
* 17406 20166: contig of 2761 bp in length
* 20167 20266: gap of 100 bp
* 20267 23474: contig of 3208 bp in length
* 23475 23574: gap of 100 bp
* 23575 26403: contig of 2829 bp in length
* 26404 26503: gap of 100 bp
* 26504 29899: contig of 3396 bp in length
* 29900 29999: gap of 100 bp
* 30000 34929: contig of 4930 bp in length
* 34930 35029: gap of 100 bp
* 35030 41463: contig of 6434 bp in length
* 41464 41563: gap of 100 bp
* 41564 47890: contig of 6327 bp in length
* 47891 47990: gap of 100 bp
* 47991 56092: contig of 8102 bp in length
* 56093 56192: gap of 100 bp
* 56193 63063: contig of 6871 bp in length
* 63064 70819: contig of 7556 bp in length
* 70820 70920: gap of 100 bp
* 70921 82582: contig of 11563 bp in length
* 82583 82645: gap of 100 bp
* 82646 98745: contig of 16063 bp in length
* 98746 117782: contig of 15037 bp in length
* 117783 149425: gap of 100 bp
* 149426 149425: contig of 31543 bp in length.
Location/Qualifiers
1. 149425
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"


```
/clone="RP11-36003"
/clone.lib="RPC1-11 Human Male BAC"
1..8144
/misc_feature /note="assembly_fragment"
vector_end:SP6
vector_side:left"
8245..9350
/misc_feature /note="assembly_fragment"
9451..11335
/misc_feature /note="assembly_fragment"
11436..14325
/misc_feature /note="assembly_fragment"
14426..17305
/misc_feature /note="assembly_fragment"
17406..20166
/misc_feature /note="assembly_fragment"
20267..22474
/misc_feature /note="assembly_fragment"
23575..26403
/misc_feature /note="assembly_fragment"
26504..29899
/misc_feature /note="assembly_fragment"
30000..34929
/misc_feature /note="assembly_fragment"
35030..41463
/misc_feature /note="assembly_fragment"
41564..47890
/misc_feature /note="assembly_fragment"
47991..56092
/misc_feature /note="assembly_fragment"
56193..63063
/misc_feature /note="assembly_fragment"
63164..70819
/misc_feature /note="assembly_fragment"
70920..82482
/misc_feature /note="assembly_fragment"
82583..98645
/misc_feature /note="assembly_fragment"
98746..117782
/misc_feature /note="assembly_fragment"
117883..149425
/misc_feature /note="assembly_fragment"
vector_end:T7
vector_side:right"
BASE COUNT 36100 a 37729 c 37628 g 36165 t 1803 others
ORIGIN
Alignment Scores:
Pred. No.: 514 Length: 149425
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.81% Indels: 0
Gaps: 0
DB: 2
US-09-872-523-5 (1-498) x AC013320 (1-149425)
QY 351 LeuG1uAspSerPheProIleHisIle 359
Db 20518 TTGGAAGACAGTTTCTTTCATTCACATC 20492
RESULT 13
AC008856/c 154229 bp DNA linear PRI 19-JUL-2001
LOCUS AC008856 Homo sapiens chromosome 5 clone CTD-2178N20, complete sequence.
ACCESSION AC008856
VERSION AC008856.6 GI:14916142
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 154229)
```

```
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 154229)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 154229)
REFERENCE DOE Joint Genome Institute and Stanford Human Genome Center.
AUTHORS Direct Submission
TITLE Submitted (19-JUL-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Jul 19, 2001 this sequence version replaced gi:13676976.
Draft Sequence. Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 100% of Sequence;
Estimated Total Number of Errors is 0.
FEATURES
source 1..154229
location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2178N20"
BASE COUNT 49945 a 27443 c 26935 g 49906 t
ORIGIN
Alignment Scores:
Pred. No.: 529 Length: 154229
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.81% Indels: 0
Gaps: 0
DB: 9
US-09-872-523-5 (1-498) x AC008856 (1-154229)
QY 359 ILEASAsnThrPheMetPheProVal 367
Db 18276 ATAAACACACCTTATGTCCACATC 18250
RESULT 14
AC016409 157102 bp DNA linear HTG 03-MAR-2000
AC016409 Homo sapiens clone RP11-114022, WORKING DRAFT SEQUENCE, 14
LOCUS AC016409
DEFINITION Unordered pieces.
ACCESSION AC016409
VERSION AC016409.4 GI:7144774
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 157102)
REFERENCE Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens, clone RP11-114022
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 157102)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Bouhgalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooper,P., Dearthiano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
```

TITLE
JOURNAL
COMMENT

Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (26-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 3, 2000 this sequence version replaced gi:5980376.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIMR

Web Site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: 114.O.22

Center clone name: 114.O.22

Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 150953 bases at least Q40

Consensus quality: 153775 bases at least Q30

Consensus quality: 154865 bases at least Q20

Insert size: 15200; agarose-1p

Insert size: 155802; sum-of-contigs

Quality coverage: 5.0 in Q20 bases; agarose-1p

Quality coverage: 4.9 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

1      872: contig of 872 bp in length
*      873      972: gap of 100 bp
*      973      4028: contig of 3056 bp in length
*      4029      4128: gap of 100 bp
*      4129      8982: contig of 4854 bp in length
*      8983      9082: gap of 100 bp
*      9083      14834: contig of 5752 bp in length
*      14835      14934: gap of 100 bp
*      14935      21185: contig of 6251 bp in length
*      21186      21285: gap of 100 bp
*      21286      26083: contig of 4798 bp in length
*      26084      26183: gap of 100 bp
*      26184      33624: contig of 7441 bp in length
*      33625      33724: gap of 100 bp
*      33725      40338: contig of 6514 bp in length
*      40339      40338: gap of 100 bp
*      40339      55110: contig of 14772 bp in length
*      55111      55210: gap of 100 bp
*      55211      68411: contig of 13201 bp in length
*      68412      68511: gap of 100 bp
*      68512      82398: contig of 13887 bp in length
*      82399      82498: gap of 100 bp
*      82499      99701: contig of 17203 bp in length
*      99702      99801: gap of 100 bp
*      99802      124153: contig of 24352 bp in length
*      124154      124253: gap of 100 bp
*      124254      157102: contig of 32849 bp in length.

```

FEATURES

Source

```

1. 157102
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Rp11-114022"
/clone_1lb="RPC1-11 Human Male BAC"
1. 872
/misc_feature
/note="assembly_fragment"

```

```

clone_end:77
vector_side:left"
misc_feature
973..4028
/note="assembly_fragment"
misc_feature
4129..8982
/note="assembly_fragment"
misc_feature
9083..14834
/note="assembly_fragment"
misc_feature
14935..21185
/note="assembly_fragment"
misc_feature
21286..26083
/note="assembly_fragment"
misc_feature
26184..33624
/note="assembly_fragment"
misc_feature
33725..40338
/note="assembly_fragment"
misc_feature
40339..55110
/note="assembly_fragment"
misc_feature
55211..68411
/note="assembly_fragment"
misc_feature
68512..82398
/note="assembly_fragment"
misc_feature
82499..99701
/note="assembly_fragment"
misc_feature
99802..124153
/note="assembly_fragment"
misc_feature
124254..157102
/note="assembly_fragment"
BASE COUNT 50151 a 28016 c 28164 g 49471 t 1300 others
ORIGIN

```

```

Alignment Scores:
Pred. No.: 537 Length: 157102
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.81% Indels: 0
DB: 2 Gaps: 0

```

US-09-872-523-5 (1-498) x AC016409 (1-157102)

OY 359 llaasnastrpmpetpheproval 367

Db 137976 ATAAACAACCTTATGTCCAGTC 138002

RESULT 15

BX323824/C

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 162649)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

```

Direct Submission
Submitted (19-APR-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zf1sh-help@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Apr 20, 2003 this sequence version replaced gi:29825483.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zf1sh-help@sanger.ac.uk
----- Project Information

```

```
Center project name: ZK37F18
----- Summary Statistics -----
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 149039 bases at least Q40
Consensus quality: 152743 bases at least Q30
Consensus quality: 156027 bases at least Q20
Insert size: 160649; sum-of-contigs
Insert size: 181199; 3.2% error; agarose-fp
Quality coverage: 4.86x in Q20 bases; sum-of-contigs quality
coverage: 4.48x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 21 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
6362: contig of 6362 bp in length
6363
6463
14590: contig of 8128 bp in length
14591
14691
14690: gap of 100 bp
34438: contig of 19748 bp in length
34439
34539
46632: contig of 12094 bp in length
46633
46732: gap of 100 bp
53161: contig of 6429 bp in length
53162
53261: gap of 100 bp
53262
59633: contig of 6372 bp in length
59634
59733: gap of 100 bp
59734
74713: contig of 14980 bp in length
74714
74813: gap of 100 bp
77867: contig of 3054 bp in length
77868
77967: gap of 100 bp
83463: contig of 5496 bp in length
83464
83563: gap of 100 bp
89435: contig of 5872 bp in length
89436
89535: gap of 100 bp
94337: contig of 4802 bp in length
94338
94437: gap of 100 bp
117110: contig of 22673 bp in length
117111
117210: gap of 100 bp
125419: contig of 8209 bp in length
125420
125519: gap of 100 bp
129093: contig of 3574 bp in length
129094
129193: gap of 100 bp
131989: contig of 2796 bp in length
131990
132089: gap of 100 bp
137224: contig of 5135 bp in length
137225
137324: gap of 100 bp
142286: contig of 4962 bp in length
142287
142386: gap of 100 bp
153725: contig of 11339 bp in length
153726
153825: gap of 100 bp
155913: contig of 2088 bp in length
155914
156013: gap of 100 bp
160055: contig of 4042 bp in length
160056
160155: gap of 100 bp
162649: contig of 2494 bp in length.
162649
Location/Qualifiers
1. .162649
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone_lib="DKEY-37F18"
/clone_lib="DantioKey"
1. .6362
/note="assembly_fragment:01786
fragment_chain:1"
6463. .14590
/note="assembly_fragment:00991
```

```
fragment_chain:1"
14691. .34438
/note="assembly_fragment:01163
fragment_chain:1"
34539. .46632
/note="assembly_fragment:02246
fragment_chain:1"
46733. .53161
/note="assembly_fragment:00743
fragment_chain:1"
53262. .59633
/note="assembly_fragment:02049
fragment_chain:1"
59734. .74713
/note="assembly_fragment:00222
fragment_chain:1"
74814. .77867
/note="assembly_fragment:02117
fragment_chain:1"
77968. .83463
/note="assembly_fragment:00362
fragment_chain:1"
83564. .89435
/note="assembly_fragment:00819
fragment_chain:1"
89536. .94337
/note="assembly_fragment:00503
fragment_chain:1"
94438. .117110
/note="assembly_fragment:00155
fragment_chain:2"
117211. .125419
/note="assembly_fragment:02285
fragment_chain:2"
125520. .129093
/note="assembly_fragment:01542
fragment_chain:2"
129194. .131989
/note="assembly_fragment:02189
fragment_chain:2"
132090. .137224
/note="assembly_fragment:00132
fragment_chain:3"
137325. .142286
/note="assembly_fragment:02023
fragment_chain:3"
142387. .153725
/note="assembly_fragment:01215
fragment_chain:4"
153826. .155913
/note="assembly_fragment:02136
fragment_chain:4"
156014. .160055
/note="assembly_fragment:00247"
160156. .162649
/note="assembly_fragment:01452"
BASE COUNT 47899 a 29936 c 31911 g 50890 t 2013 others
ORIGIN
Alignment Scores:
Pred. No.: 554
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.81%
DB: 2
US-09-872-523-5 (1-498) x BX323824 (1-162649)
QY 395 LysGluSerAlaGluThrLeuProLeu 403
Db 107178 AAAGAGCTCGAGAAACACTTCCTTTC 107152
```

Search completed: September 2, 2003, 22:38:48
Job time : 6693 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2003, 13:45:39 ; Search time 5682 Seconds

(without alignments)
10778.196 Million cell updates/sec

Title: US-09-872-523-6

Sequence: 1 atgtcgaattctgaaat.....cgaaaagtacacattga 1497

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*
1: gb_da:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pln:*
35: em_hlg_rtd:*
36: em_hlg_mam:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_hlgo_hum:*
40: em_hlgo_mus:*
41: em_hlgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	1103	73.7	30911	3 CER06C7	271266.1 Caenorhabd
c 2	422	28.2	99298	3 AC090416	AC090416 Caenorhab
c 3	139.2	9.3	43096	3 AC024792	AC024792 Caenorhab
c 4	128.8	8.6	31267	2 AC006785	AC006785 Caenorhab
5	60.6	4.0	2878	10 BC020018	BC020018 Mus muscu
6	59	3.9	164681	2 AC101996	AC101996 Mus muscu
7	59	3.9	215435	10 AL662838	AL662838 Mouse DNA
8	58.8	3.9	2960	6 AF285577	AF285577 Mus muscu
9	58.8	3.9	2969	6 AX247680	AX247680 Sequence
10	58.8	3.9	163349	10 AL672233	AL672233 Mouse DNA
11	58.8	3.9	230185	2 AC021579	AC021579 Mus muscu
12	57.4	3.8	3590	6 AX046388	AX046388 Sequence
13	55.8	3.7	2564	9 AK000062	AK000062 Homo sapi
14	55.8	3.7	128328	9 AC005839	AC005839 Homo sapi
15	55.8	3.7	199301	2 BX088589	BX088589 Danio rer
16	55.4	3.7	251144	2 AC109565	AC109565 Rattus no
17	55.4	3.7	258549	2 AC133113	AC133113 Rattus no
18	55.2	3.7	2425	9 AK092756	AK092756 Homo sapi
19	55.2	3.7	3255	9 AK094137	AK094137 Homo sapi
20	55.2	3.7	4101	9 AB058701	AB058701 Homo sapi
21	55	3.7	3385	10 BC030864	BC030864 Mus muscu
22	55	3.7	3399	10 AY237001	AY237001 Mus muscu
23	55	3.7	3725	10 AY237002	AY237002 Mus muscu
24	54.2	3.6	2099	9 AK093255	AK093255 Homo sapi
25	54.2	3.6	2845	9 AK091075	AK091075 Homo sapi
26	54.2	3.6	4534	9 BC039316	BC039316 Homo sapi
27	53.8	3.6	193660	2 AC102262	AC102262 Mus muscu
28	53.4	3.6	1035	6 AX576074	AX576074 Sequence
29	53.4	3.6	2105	9 AB072769	AB072769 Macaca fa
30	53.4	3.6	3250	9 AF149045	AF149045 Homo sapi
31	53.4	3.6	3277	9 BC021252	BC021252 Homo sapi
32	53.4	3.6	3315	9 BC009752	BC009752 Homo sapi
33	53.4	3.6	3321	9 AF149046	AF149046 Homo sapi
34	53.4	3.6	136649	9 HSD635E8	AL110502 Human DNA
35	52.4	3.5	6480	5 FRNINE2	AF146688 Fugu rubr
36	51.8	3.5	3105	6 BD158238	BD158238 Primer fo
37	51.8	3.5	3105	9 AK023124	AK023124 Homo sapi
38	51.4	3.4	3096	9 BC051913	BC051913 Homo sapi
39	51.4	3.4	4130	9 HSA18004	Y18004 Homo sapien
40	51.4	3.4	47167	9 HS757P12	AL031007 Human DNA
41	50.6	3.4	2432	9 AK097052	AK097052 Homo sapi
42	50.6	3.4	2555	6 AX078286	AX078286 Sequence
43	50.6	3.4	3071	9 HSM801538	AL136564 Homo sapi
44	50.6	3.4	3188	6 AX714233	AX714233 Sequence
45	50.6	3.4	3188	9 AK056443	AK056443 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS CER06C7 30911 bp DNA linear 271266.1-271266.1
DEFINITION Caenorhabditis elegans cosmid R06C7, complete sequence.
ACCESSION 271266
VERSION 271266.1 GI:1279324
KEYWORDS HTG; Adenylosuccinate lyase; Myosin heavy chain; Rat TOAD-64
protein like; Serine/threonine-protein kinase; Zinc finger protein.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1
AUTHORS none.
TITLE genome sequence of the nematode C. elegans: a platform for

JOURNAL	MEDLINE	PUBMED	REMARK	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Investigating biology. The C. elegans Sequencing Consortium	99069613	9851916	The C. elegans Sequencing Consortium.	2 (bases 1 to 30911)	Gardner, A.E.	Direct Submission	Submitted (19-APR-1996) Nematode Sequencing Project, Sanger Institute, Hinxton, Cambridge CB10 1SA, England and Department of Genetics, Washington University, St. Louis, MO 63110, USA. E-mail: jess@anger.ac.uk or rwhematode.wustl.edu	Coding sequences below are predicted from computer analysis, using predictions from Genefinder (P. Green, U. Washington), and other available information.
								Current sequence finishing criteria for the C. elegans genome sequencing consortium are that all bases are either sequenced unambiguously on both strands, or on a single strand with both a dye primer and dye terminator reaction, from distinct subclones. Exceptions are indicated by an explicit note.
								IMPORTANT: This sequence is not the entire insert of clone R06C7. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.
								The true left end of clone R06C7 is at 1 in this sequence. The true right end of clone R06C7 is at 8718 in
								sequence Z71261.
								The true left end of clone F21C3 is at 30808 in this sequence. The start of this sequence (1..106) overlaps with the end of sequence Z98261.
								The end of this sequence (30808..30911) overlaps with the start of sequence Z71261.
								For a graphical representation of this sequence and its analysis see: - http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?name=R06C7
								IMPORTANT: This sequence is NOT necessarily the entire insert of the specified clone. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.
								Location/Qualifiers
								1..30911
								/organism="Caenorhabditis elegans"
								/mol_type="genomic DNA"
								/strain="Bristol N2"
								/db_xref="taxon:6239"
								/chromosome="1"
								/clone="R06C7"
								join(complement(588..775),complement(395..539),complement(169..295),complement(298261.1:204..368),complement(298261.1:105..155))
								complement(271258.1:40294..40359),
								complement(271258.1:40101..40244),
								complement(271258.1:39857..39994),
								complement(271258.1:39258..39701),
								complement(271258.1:38262..39212),
								complement(271258.1:38105..38207),
								complement(271258.1:37940..38057),
								complement(271258.1:37773..37895))
								/gene="C01H6.9"
								join(complement(588..775),complement(395..539),complement(169..295),complement(298261.1:204..368),complement(298261.1:105..155))
								complement(271258.1:40294..40359),
								complement(271258.1:40101..40244),
								complement(271258.1:39857..39994),
								complement(271258.1:39258..39701),
								complement(271258.1:38262..39212),
								complement(271258.1:38105..38207),
								complement(271258.1:37940..38057),
								complement(271258.1:37773..37895))
								/gene="C01H6.9"
								/standard_name="C01H6.9"
								/note="contains similarity to Pfam domain: PF00069

(Eukaryotic protein kinase domain) score=37.5,
E-value=8.2e-10, N=2
cDNA EST yk42h5.3 comes from this gene"
cDNA EST yk42h5.3 comes from this gene"
/codon_start=1
/product="Hypothetical protein C01h6.9"
/protein_id="CAA95847.1"
/db_xref="GI:3878842"
/db_xref="GOA:017985"
/db_xref="SPTRMBL:O17985"
/translation="MPPKRIKVPISPGKARKNNFARRNRKQJSSVITTCPIDPEYE
EKNALFSEFNDNEPFKSKKNKRPPEPSPVSDVEISDEDDAQFVPPDDA
DEEEKIKTKKEAFSLDDVSLFENNNDGRTVAIEENKHI, IHRKPEAPRPPEAKID
EKKKMKRQKRVKQVPEMNSKANQVOCGTANVCSGSLILADHRONKTKASLAAPL
RKRGCPQKSKMIVPAKHQCCODERLIAVSTPTRVKTKKOLLIDTIGSLVFSKTONA
STPIEOKITGRPKSPVVDVLSLSCENCITTKREEDIAKNSANSODKETTINPIEE
ESGVRGSHSLAGS IQQPOQDLEETAISNRNLSHTSLKTPLESRNESHQNSKIMTQ
VSDSEVDENVTFRNKAETSVTEILHEDHETQPODRLNEESKSRKQRTKSPETIQ
FLEDMNGEGRGTQYSESRAESNRITPTGMTINHEDSLITPLEDLTSEIDSPMKQ
LHVQGEKERTMDSLPKSLDGRVYKLGEGAVGFSITWQKPAIKIVPEKQGC
NQVQRETHSEMQSDVYLPETVYKELISALDEDPANSTPTFTIMISAEMVGGKYP
KILLSMWSAEALKESENRDPVYSIDQNLIFLVSANGALAEVLSEBNLSIT
HOLIVSNMAEALKESENRDLHGNVLIDRNGYKELYVHVGQKVLSTHGIKVNLID
FLTSRSKATTVYMDLENDPAIFEGODDPQDFEYVEMKRNCKNSMKKRSRRNLMI
VIANLIDITKICPKGLLEKRMELKVLFDRAEFGSGCESLITNEEFSDFEYEGIG
MSTTRQ"
complement(join(1315. 1480,1528. 1715,1747. 1783,
1834. 2156))
/gene="R06C7.2"
complement(join(1315. 1480,1528. 1715,1747. 1783,
1834. 2156))
/gene="R06C7.2"
/standard_name="R06C7.2"
/codon_start=1
/product="Hypothetical protein R06C7.2"
/protein_id="CAA95841.1"
/db_xref="GI:3878836"
/db_xref="SPTRMBL:O21772"
/translation="MFSAVHPDSRCKKRYRSQTSNONYWFLYSEPNLGRKITG
NFIHLHAERKVLQSVMLKDQYAKTEKRTQELTQCSLANFEENRKKLODTREN
KPRNLINSLTELKLVSSPENLPIQLIFLISENDELIERFHCATFKLSPLFLE
FPPRVVANAESAQSGKSNLMNEALITADAEVSEAPKARMQIDSIKXKSGFLIDRR
VNSALLIOSNCISFNSNF"
join(3063. 3227,3274. 3854,3899. 4631,4680. 4855,
4899. 5074,5117. 5463,5510. 5749,5790. 6097,6143. 6254)
/gene="R06C7.1"
join(3063. 3227,3274. 3854,3899. 4631,4680. 4855,
4899. 5074,5117. 5463,5510. 5749,5790. 6097,6143. 6254)
/gene="R06C7.1"
/standard_name="R06C7.1"
/note="contains similarity to Pfam domain: PFO2170 (ZAP
domain), score=151.7, E-value=4.2e-42, N=1; PFO2171 (P1w1
domain), score=456.7, E-value=6.3e-134, N=1
cDNA EST yk31a12.5 comes from this gene
cDNA EST yk21g1.3 comes from this gene
cDNA EST yk25b2.3 comes from this gene
cDNA EST yk31a12.3 comes from this gene
cDNA EST yk31a12.3 comes from this gene
cDNA EST yk36g4.3 comes from this gene
cDNA EST yk36g4.3 comes from this gene
cDNA EST yk31a4.3 comes from this gene
cDNA EST yk21g1.5 comes from this gene
cDNA EST yk25b2.5 comes from this gene
cDNA EST yk34a6.5 comes from this gene
cDNA EST yk36g4.5 comes from this gene
cDNA EST yk31a4.5 comes from this gene
cDNA EST yk31h11.5 comes from this gene
cDNA EST yk125b7.5 comes from this gene
cDNA EST yk125b7.5 comes from this gene
cDNA EST yk125b7.3 comes from this gene
cDNA EST yk236e4.3 comes from this gene
cDNA EST yk295d11.3 comes from this gene
cDNA EST yk148h1.3 comes from this gene
cDNA EST yk367e6.3 comes from this gene

```

cDNA EST yk236e4.5 comes from this gene
cDNA EST yk295d11.5 comes from this gene
cDNA EST yk348b1.5 comes from this gene
cDNA EST yk367e6.5 comes from this gene
cDNA EST yk513b8.3 comes from this gene
cDNA EST yk669b4.3 comes from this gene
cDNA EST yk671a8.3 comes from this gene
cDNA EST yk250f12.5 comes from this gene
cDNA EST yk513b4.5 comes from this gene
cDNA EST yk669b4.5 comes from this gene
cDNA EST yk671a8.5 comes from this gene
cDNA EST yk775g12.3 comes from this gene
cDNA EST yk775g12.5 comes from this gene
/codon_start=1
/product="Hypothetical protein R06C7.1"
/protein_id="CA95839.1"
/db_xref="GI:3878834"
/db_xref="SPTREMBL:Q21770"
/translacion="MSPHPQPHPPMPMPVPVAPPGAMTPMPVPVADAKLHQSTGN
DACIKRLQQLNVEDGAKMYMKPTPEKMGKRPVDIQIQTAVFEGIEVTKETTVHREYHAKA
DLSTKEVETTKKGEDEVYQDRDKCNLEPLAVENKPEPFMKDNOIYVGSQSL
YTVNLSELDANGTSKVRQINGADNDIKTLCISLEIYARPNSTITLSENG
KTRADONIEVNNREYTOFELALNHCVRBTNRGCFEHRKVFPLNTEGFDQDRCV
DVGDKQLYFGLKTIQIEGPGYGRGNNSLVLDGKKAHFHQVYIQKLFIDTGD
PSGGLNNMTRKKAAYIKGLDCYSTYNNRRLRIEGLIFHESATKTFEELPDGTCST
AEYVADKYKISLQYNNANLVCKDRGNNTYPAELMVSNNORATLPOQGNOSOKT
KECAVLDVROHMITGKNAVNITLENELVALGIKYKSPPLWVAQAEIDGKELVGR
SVASDMKMRAPPGMPYKRPATVPDIAAAYVGNRGCFSDVNOIYGMPTDSCKKKG
MYIKPREGTLYSTEKIMTOLKRYAKSKCYVLMITDDALVHLKQKALEQRTMYLV
QDMKISKANAIVKDGKRLLENLINTKNVGLGNTVSPAKSMDEBOLIIGVSA
PPAGTKYMDMKNHNLNPOIIGFASNAVAHNEFGVDPLASGDVTMASIEDVONSID
LPEKNRAALPKRIIYRSGASEGSHASILAIEPLAIIHGYSKEIKLIEIVYTKH
SYEFRRDOLRSGSKATEMINIPGVIDNATNPACQFPLNGHTTLOGCTAKTPLYTVL
ADDCKAPMDRLSELTFLCHHIOIVLSSTSIPTPIYANEXARKGRDLNMLELTKGI
EAKESGGERLKEITKEIGYKOTDLNQRVNA"
/gene="dhp-1"
/complement(6739..11408)
cDS
complement(join(6739..6889,7899..7963,8094..8340,
8392..8457,8905..9182,9550..9761,9840..10017,10603..10783,
11317..11408))
/gene="dhp-1"
Query Match 73.7%; Score 1103; DB 3; Length 30911;
Best Local Similarity 81.3%; Pred. No. 1.5e-240;
Matches 1497; Conservative 0; Mismatches 0; Indels 344; Gaps 5;
QY 1 ATGTCTGAATTTCTGAAAAATTCAGAGCTAACAAAAAATCGGACAGAAAACCTCGATTAAG 60
Db 18471 ATGTCTGAATTTCTGAAAAATTCAGAGCTAACAAAAAATCGGACAGAAAACCTCGATTAAG 18412
QY 61 ACCTACTTGTGGGAATCCTATTACATCAGTTCGAGAGAGAAAAAATCTTTCATTCGA 120
Db 18411 ACCTACTTGTGGGAATCCTATTACATCAGTTCGAGAGAGAAAAAATCTTTCATTCGA 18352
QY 121 GTTGAAGCATTCATCGTAACTTACAGTTAATTTTAAAGAAATGCGTGAAGGAAGATT 180
Db 18351 GTTGAAGCATTCATCGTAACTTACAGTTAATTTTAAAGAAATGCGTGAAGGAAGATT 18292
QY 181 ATC-----TTGGAACA 192
Db 18291 ATGCTGAGTTCATATTGTTGCTAAATCGGTTTAAATATCAATTTTGTAGTTGGAACA 18232
QY 193 GTGTCCATGATTATGACAGAACTGCGATTTCGATTCAGTCAAGATGTTTGGACGAATT 252
Db 18231 GTGTCCATGATTATGACAGAACTGCGATTTCGATTCAGTCAAGATGTTTGGACGAATT 18172
QY 253 GAAAAAGTTTGGCGATACAGAGTTCTGGCTCAGTTTATCGAGACTGACAGAAATTTGG 312
Db 18171 GAAAAAGTTTGGCGATACAGAGTTCTGGCTCAGTTTATCGAGACTGACAGAAATTTGG 18112
QY 313 CTCGAATATTTATCGAGCATATGTTGGTTTGGCAAA----- 350
Db 18111 CTCGAATATTTATCGAGCATATGTTGGTTTGGCAAAATTAAGTTGAGACCTCAGCTCTT 18052
```

```

QY 351 ----- 350
Db 18051 TCTACTATTTCTAAATTAATATGTTCTGTTACATAAAATTCYAGACAAATGCTATTA 17992
QY 351 -----CGCCGAATGAGTAT 366
Db 17991 AAATCTGAAACATTGTATATAATGTAATAATTTGAACTTTCACGGCGCGCAATGAGTAT 17932
QY 367 CCCAATATGATATAAATTTGATATAGCTCCGCCCTTGCAATCAACGAAGAAATACCAAAAT 426
Db 17931 CCCAATATGATATAAATTTGATATAGCTCCGCCCTTGCAATCAACGAAGAAATACCAAAAT 17872
QY 427 GATATGTAATTTGTAAT----- 447
Db 17871 GATATGTAATTTGTAATGTAATGTTGTTTTTTCCGAATTTAGTTAATATCATCT 17812
QY 448 -----AATTGATATGATGCGCAAAATCGTCGGCAAACTTCGCTGCTCCAAAT 496
Db 17811 CACAACCTTCAGAAATTTGATATGATGCGCAAAATCGTCGGCAAACTTCGCTGCTCCAAAT 17752
QY 497 TCGATGAAGGGAAGGCTCTCTTAAGCAAGCATCGTTTCAAAGTTGACACACGCTTGAC 556
Db 17751 TCGATGAAGGGAAGGCTCTCTTAAGCAAGCATCGTTTCAAAGTTGACACACGCTTGAC 17692
QY 557 TATTAAATTTATTCGAATTTCTGAAATPACGCGTACGCGCAATTCAGAAATATGTGGAC 616
Db 17691 TATTAAATTTATTCGAATTTCTGAAATPACGCGTACGCGCAATTCAGAAATATGTGGAC 17632
QY 617 GACGAATGAATATATCTATCAAAAGAAAGACTTCCCGAATCGCTTCAGATCAGATG 676
Db 17631 GACGAATGAATATATCTATCAAAAGAAAGACTTCCCGAATCGCTTCAGATCAGATG 17572
QY 677 ACGACACAAAGTCTTAAAGCTCTGATCTCAATTTTGGATAGACAGAGGAAGCTTCTCA 736
Db 17571 ACGACACAAAGTCTTAAAGCTCTGATCTCAATTTTGGATAGACAGAGGAAGCTTCTCA 17512
QY 737 TATTTCTGTTGATTTGACAGACATCAATGATATCAATTAATGCGAAAAAGGAATATA 796
Db 17511 TATTTCTGTTGATTTGACAGACATCAATGATATCAATTAATGCGAAAAAGGAATATA 17452
QY 797 TTGAGCACACAAATTAATTTGCTCAAGCAATTAATAAATGAGAAATCCAAATATGACT 856
Db 17451 TTGAGCACACAAATTAATTTGCTCAAGCAATTAATAAATGAGAAATCCAAATATGACT 17392
QY 857 CAGACGACGTCACATTTGATCAATTAACAAAAGATCCAAATGATCCATGATTTGGAGAA 916
Db 17391 CAGACGACGTCACATTTGATCAATTAACAAAAGATCCAAATGATCCATGATTTGGAGAA 17332
QY 917 AAGTTAAGTTTGGACAAAGTTTGAAGTCATCGACCCCTTGAGTCAATTAATTAAC 976
Db 17331 AAGTTAAGTTTGGACAAAGTTTGAAGTCATCGACCCCTTGAGTCAATTAATTAAC 17272
QY 977 TCACGTCGCTTCGATTTCTCAAAATTTGCAAAACCTGAAGATATCTTATGTGGAAATG 1036
Db 17271 TCACGTCGCTTCGATTTCTCAAAATTTGCAAAACCTGAAGATATCTTATGTGGAAATG 17212
QY 1037 ATGCTCCAGATGCATTTGAAGACGTTTCTATTCATATCAATTAATTAATTAATTTGCC 1096
Db 17211 ATGCTCCAGATGCATTTGAAGACGTTTCTATTCATATCAATTAATTAATTAATTTGCC 17152
QY 1097 CAGTTGTTATGCGGAAAGATATTAATTTGAACTTGTCCGCCGATGAGTTCAAAGAA 1156
Db 17151 CAGTTGTTATGCGGAAAGATATTAATTTGAACTTGTCCGCCGATGAGTTCAAAGAA 17092
QY 1157 CATTCAGATGGATGAATTAATTTGAGAAAGAAATTCGAGAAACCTACCGCTTACCTGT 1216
Db 17091 CATTCAGATGGATGAATTAATTTGAGAAAGAAATTCGAGAAACCTACCGCTTACCTGT 17032
QY 1217 TCAAGCAATATGCTTCCCAAGAGAGATTAACAAATTTAAGTATTTCTGATTTCCAAAC 1276
Db 17031 TCAAGCAATATGCTTCCCAAGAGAGATTAACAAATTTAAGTATTTCTGATTTCCAAAC 16972
```

	QY	1277	G	-----	1277	1	-----	1277
Db	16971	GGGTTGTTTAA	TCGTTGAGATGTTTCACTAATTAAGTATTCATTAATGTTTCTT		16912			
QY	1278	-----	GGTAGGACTACGCCCTTGAAGCTGCCTGACATGTCGTAACATCAGTTATTTGTC		1330			
Db	16911	GTTTAAAGTAGGAC	TACACGCCCTTGGAAGCTCTGCACATGTCGTAACATCAGTTATTTGTC		16852			
QY	1331	CAGCTACAGTGA	AAATCAGTTTCATGGAAGACTGATTAATGTCATTTTCGACGCTGGGATG		1390			
Db	16851	CACCTACAGGA	AAATCAGTTTCATGGAAGACTGATTAATGTCATTTTCGACGCTGGGATG		16792			
QY	1391	AGCAATTTGAT	GAACTGTATGATGTGGA-----		1418			
Db	16791	AAGCAATTTGAT	GAACTGTATGATGTGGAAGACTGATTAATCATTGACCGAAGACATTTTTC		16733			
QY	1419	-----	CTCCCATGATATTTCTACCGATAGGATGCTGTGAAGCGC		1456			
Db	16731	AATGAAATTC	TATTCATATTCAGTCCCATATATTTCTACCGATAGGATGCTGTGAAGCGC		16672			
QY	1457	ACAGTTATG	TTCTACACACTCCGAAAAGTACAACTTTGA		1497			
Db	16671	ACAGTTATG	TTCTACACACTCCGAAAAGTACAACTTTGA		16631			
RESULT 2								
AC090416/c								
LOCUS	AC090416		99298 bp	DNA	linear	INV 17-FEB-2001		
DEFINITION	Caenorhabditis briggsae		cosmid CB042D07,		complete sequence.			
ACCESSION	AC090416							
VERSION	AC090416.1		GI:12958081					
KEYWORDS	HTG.							
SOURCE	Caenorhabditis briggsae							
ORGANISM	Caenorhabditis briggsae							
REFERENCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea; Rhabditidae; Pelodermidae; Caenorhabditis.							
AUTHORS	1 (bases 1 to 99298)							
TITLE	Washington University Genome Sequencing Center.							
JOURNAL	The C. briggsae Genome Sequencing Project							
REFERENCE	Unpublished							
AUTHORS	2 (bases 1 to 99298)							
TITLE	Waterston, R.							
JOURNAL	Direct Submission							
COMMENT	Submitted (17-FEB-2001) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA							
	Submitted by:							
	Genome Sequencing Center							
	Department of Genetics, Washington University,							
	St. Louis, MO 63110, USA							
	e-mail: jspiehl@watson.wustl.edu							

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

FEATURES	Location/Qualifiers
source	1. .99298

| complement(67280. .67362) | trna |

BASE COUNT	29971 a	19350 c	19512 g	30465 t
------------	---------	---------	---------	---------

Query Match	28.2%;	Score 422;	DB 3;	Length 99298;
Best Local Similarity	62.5%;	Pred. No. 1.3e-85;		
Matches 689;	Conservative	0;	Mismatches 390;	Indels 23;
				Gaps 1;

OY	356	CAATGAGTATCCCAATATGGAATAATGTATATGCTCGCGCGGTTCATCAACGAAG	415
Db	18471	CGATGATCCCAATGTGATCGAAGACCAATATCTATGTCTCACACCAAAATGTTATGAGCC	18412
OY	416	AATACCAAAATGATATGTAATTAATGTAAATAATTCGATTTGATGGCGAAATCGTCGCC	475
Db	18411	TGTACCAACGATGATTTGGATGACTTTTATGTACAAAAACAATGATGATGAATGGATGG	18352
OY	476	AAACTTCGCTGTCTCCAAAAATTCGATGAAGGAGAGCTCTCTTAAGCAACATGTTTTCA	535
Db	18351	AATTTGAGTCTGTCTCCCGGTTACGATGAGAAATAAAGAGAGCTCTCAAGTCGTGTTCC	18292
OY	536	AAATTGGACACACTCTTGAACATTAATTAATTTCCAAATCTCTCAAAATACGCGTAGGC	595
Db	18291	GCGTTGGACAGCGCTGGAGCTTCTCACTACGCAACTTCTACAGAAAATAGCTGTCGAC	18232
OY	596	GAATTCAAGAAATATGTGACGACGACGAATGATATCTATCAACAAAGAAAGACTTTCCG	655
Db	18231	GTGTTGAGAAAGTCGTGTGGTCGTGGATGACAGTTCTGTGTCACCAAAAGACTATTCAG	18172
OY	656	AATGCTTCCAGATTCGACAGATGACGACAGCAAGTCTTTAGCTCTGATCTCAATTTGGA	715
Db	18171	GAACTCTCCGAATTTCAAGAAAGACGCGCAAGTTCAAGGCTCGGAATCTGTTTTGA	18112
OY	716	TAAACGGGGGAACCTTCTCATATTTCTGTTGGATTTTCCACAGTCAGATGATATCAAC	775
Db	18111	TCGACGAGGAGACTCTTCTTGTTGTTCCAGTTGGTTTTCTGCTTTCATATCAAT	18052
OY	776	TAAATGCGAAAGAGATATATTTAGCACACACAATTAATTAATGCTCAAGCAATAAAAATG	835
Db	18051	TGCTTGCCACTGAAAGATATATACGACACACTCCAAATATCTCGCTTGATGAATG	17992
OY	836	GAGAAAATCCAAAGTATGACTGACGACGAGCTCACTTTGATCAATTAGCAAAAAGATCCA	895
Db	17991	GACAGGAGCCAGATATCATATGTAGAGACACTTGATTCACCGATTTATCAAGAGAAGATG	17932
OY	896	TTGATTCCTCATTTTGGAGAAAGTTAAGTTTGGACAAAAGTTTGACATCGACCCCT	955
Db	17931	TTGATTAAGAAATGGGAAAATATGAAATCGGACTGAAATTCGAGCTCATATGATCCAT	17872
OY	956	TGCGTCAGCAATTCATATMACCTCCAGCTCGCTTCGATTCCTCAATTTTGCAAAACCTGAG	1015
Db	17871	TGCGTCAAGATTCGAAAACCTCAAGTGGCATCGGTTATCTCGTTGTGCACACAGCAATG	17812
OY	1016	GATATCTTATTTGGGAGATGATGATGATCGTCACTGTAAAGACATTTTCTATTCAATA	1075
Db	17811	GATATATAATTTGTCGAATCGATGAGACCCGATATGAACGATGAGATTTTCCACTTCACA	17752
OY	1076	TCAATTAATACATTAATGTATCTCCAGTTGGTATATCGGAAAAGTATTAATTTGGAACCTGTTC	1135
Db	17751	TCGACAACGCTTCTATCTTCCCTGTGGATATTTGTGAGCATTAATCTGMAACTGATAA	17692
OY	1136	CGCAGATGAGTTCAAGGAACATTCAGATGGAGTGAATATCTTGGAGAAAAGAAATCTGCAG	1195
Db	17691	AGCCAGATGGGTACAAAGAAATCGTTCAAAATGGGATGAATCTGCGCGTGAGAAATGCC	17632
OY	1196	AAACCTTACCGCTTGACTTGTTCACGCCAATGCCCTTCCACAGAGATTAAGCAAAATTAA	1255
Db	17631	AAACACTTCAAATTTGAACCTTTTATAGCGCAGACGCCACTCAGGAGAACTTGAAACATTTTC	17572
OY	1256	AGGTAATTCGATTTCCAA-----ACGGGTAGAGCTTAGCGCT	1292
Db	17571	AGGTAATTCGATTCAAATAATGTGATGTAATAATTCCTAAATTTTCAGGTCGGAATGCGTCT	17512
OY	1293	TGAAGCTCTCACATGTGTGTAATAATCAAGTTATTTGTTCACGATACAGTGAATCAAGTTCA	1352
Db	17511	CGAAGCAACGATATTTGTGCGAAGAACCACTGATCAGTCCCGGAGAAATCTCTCAATTTCA	17455
OY	1353	TGGAAGATGATTAATGTCAATTTTCGACGCGCTGGATGAAGAAATTTGATGAACATGTATTA	1412
Db	17451	CGAGAGATATCAATTAATGTCAATTTTCGATGATGGAGAAAGCTTTGACGAGCTTACACA	17392
OY	1413	TGTGGACTCCCATGATTTCTTA	1434

Db 17391 TATTAAGTCAGTTTATGCGCTA 17370

RESULT 3

AC024792/c 43096 bp DNA linear INV 22-NOV-2002
DEFINITION Caenorhabditis elegans cosmid Y48G1A, complete sequence.
AC024792
AC024792.1 GI:7140348
HTG.
KEYWORDS
SOURCE
ORGANISM
Caenorhabditis elegans
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditidae; Rhabditidae; Pelodierinae; Caenorhabditis.

1 (bases 1 to 43096)
Waterston, R.
Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium
Science 282 (5396), 2012-2018 (1998)

JOURNAL
MEDLINE
PUBMED
99069613
9851916

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
2 (bases 1 to 43096)
Bemis, G., Ryan, E. and Courtney, L.
The sequence of C. elegans Y48G1A
Unpublished (2001)
3 (bases 1 to 43096)
Waterston, R.H.
Direct Submission

Submitted (01-MAR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 43096)
Waterston, R.
Direct Submission

Submitted (28-MAR-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
5 (bases 1 to 43096)
Waterston, R.
Direct Submission

Submitted (09-AUG-2001) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
6 (bases 1 to 43096)
Waterston, R.
Direct Submission

Submitted (16-NOV-2001) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
7 (bases 1 to 43096)
Waterston, R.
Direct Submission

Submitted (24-MAY-2002) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
8 (bases 1 to 43096)
Waterston, R.
Direct Submission

Submitted (28-AUG-2002) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
9 (bases 1 to 43096)
Waterston, R.
Direct Submission

Submitted (22-NOV-2002) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
Submitted by:

Genome Sequencing Center
Department of Genetics, Washington University
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1RQ, England

email: rv@nematode.wustl.edu and jess@sanger.ac.uk
NOTICE: This sequence may not be the entire insert of this clone.
It may be shorter because we only sequence overlapping sections
once, or longer because we provide a small overlap between
neighboring submissions.

This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate chemistry
or covered by high quality data (i.e., phred quality >= 30); an
attempt was made to resolve all sequencing problems, such as
compressions and repeats; all regions were covered by sequence from
more than one m13 subclone.

For a graphical representation of this cosmid sequence and its
analysis see:
<http://www.wormbase.org/db/seq/sequence?name=Y48G1A:class=Sequence>

NEIGHBORING COSMID INFORMATION

The 5' cosmid is C53D5, 1000 bp overlap; the 3' cosmid is R119, 300
bp overlap.

NOTES:

Coding sequences below are the result of integration and manual
review of the following data: computer analysis using the program
GeneFinder (P. Green and L. Hillier, personal communication), the
large scale EST projects of Yujii Kohara
(http://www.ddbj.nig.ac.jp/c-elegans/html/CE_INDEX.html) and The C.
elegans ORFome cloning project (<http://worldb.dfci.harvard.edu/>),
similarity to other proteins from Blastx analyses
(<http://blast.wustl.edu/>), sequence conservation with C. briggsae
using Jim Kent's WABA alignment program (Genome Research
10:1115-1125, 2000), individual C. elegans Genbank submissions,
and personal communications with C. elegans researchers. tRNAs
are predicted using the program tRNAscan-SE (Lowe, T.M. and
Eddy, S.R., 1997, Nucleic Acids. Res., 25, 955-964).
Location/Qualifiers

FEATURES
source

1..43096
/organism="Caenorhabditis elegans"
/mol_type="genomic DNA"
/strain="Bristol N2"
/db_xref="taxon:6239"
/chromosome="I"
/clone="Y48G1A"
complement(4415..8942)
/gene="Y48G1A.5"
complement(join(4415..4718,5328..6007,6665..7978,
8375..8683,8733..8942))
/gene="Y48G1A.5"

gene

CDs
/note="contains similarity to Pfam domains PF03378
(CAS/CSE protein, C-terminus), PF03810 (Importin-beta
N-terminal domain); coded for by the following C. elegans
cDNAs: YK75C11.3, YK75C11.5, YK85B9.5, YK19263.3,
YK19263.5, YK225612.3, YK225612.5, YK228912.3, YK228912.5,
YK381B10.5, YK4464.3, YK4464.5, YK484B6.3, YK484B6.5,
YK508B5.5, YK610A2.3, YK610A2.5, YK640B6.3, YK640B6.5,
YK642F4.3, YK642F4.5, YK650D10.3, YK650D10.5, YK732E2.3,
YK732E2.5"

/codon_start=1
/product="Hypothetical protein Y48G1A.5"
/protein_id="AAK84610.1"
/db_xref="GI:15145461"
/db_xref="wormbase:Y48G1A.5"
/translation="MEQICAAQQTLPEPAAIRKGEALRSLQSPGYIIQILLV
NEQOIIADQIRIAAVALKNEFKRMWGPAPVEVMGQEDQFRSLMEAMFYTKSNVQ
ELISNALYLAQRDEPKWPDVLPYLSKFLNADLNHLVASLASMEQIFRKRFESKS
ALMKELKKCLSTQDEPLTLLRNWVGORQDQADGADIAQMLVLLILAKVYSIC
SQEIPREYDHLKDMMPHFLHLYQIDAPQYNSNGSEPTIDELKHEICIFLYSQRV
EEIESEFPDITLAVNMLKSTGPOTRFDYPCALBEPISVYSQNYIRGHFTGSCVY
KTLAENVCVQNILRQDQWELFDEPDLDMKIDIGTVDGTRRAAIDLRGLCRFE

* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

*	1	2189:	contig of 2189 bp in length
*	2190	2202:	gap of unknown length
*	2203	5146:	contig of 2944 bp in length
*	5147	9788:	gap of unknown length
*	5160	9788:	contig of 4659 bp in length
*	9789	9801:	gap of unknown length
*	9802	15045:	contig of 5244 bp in length
*	15046	15058:	gap of unknown length
*	15059	19039:	contig of 3961 bp in length
*	19040	19052:	gap of unknown length
*	19053	25621:	contig of 6559 bp in length
*	25622	25634:	gap of unknown length
*	25635	33335:	contig of 7701 bp in length
*	33336	33348:	gap of unknown length
*	33349	41770:	contig of 8122 bp in length
*	41471	41483:	gap of unknown length
*	41484	59303:	contig of 17820 bp in length
*	59304	59316:	gap of unknown length
*	59317	74066:	contig of 14750 bp in length
*	74067	74079:	gap of unknown length
*	74080	90677:	contig of 16598 bp in length
*	90678	90690:	gap of unknown length
*	90691	113452:	contig of 22762 bp in length
*	113453	113465:	gap of unknown length
*	113466	154662:	contig of 41197 bp in length
*	154663	154675:	gap of unknown length
*	154676	31267:	contig of 157592 bp in length

BASE COUNT	a	c	g	t
ORIGIN	98825	56772	57406	99095
				169 others

Query Match	8.68;	Score 128.8;	DB 2;	Length 312267;
Best Local Similarity	53.38;	Pred. No. 6.5e-19;		
Matches 400; Conservative	0;	Mismatches 322;	Indels 28;	Gaps 5;

Qy	528	TCGTTTCAAGTGGACACGCTGTGACATTAATTAATTCACATTCCTCTGGAATACG	587
Db	112122	TCGTTTCCGGCTAATCAACGCGTCGAGTACTCACTACTTGGGCCAACAGAAATCCG	112181

QY	588	CGTGGCGCCAAATTCAGAATATATGTCGACGACGAATGAATGTATATATATCAACAAAGAAG	647
Db	112182	TGTGTCGACG-ATCTCTCAGCATCTTTGGCCCCCGTGTGATGTCATGCTACCGCTCAGGA	112240

[illegible]

OY	705 TCATATTTCGATAGACGAGGGGAACCTTCTTCATATTCCGTGGATTTTCAGCAGTCAA	764
	- - - - - - - - - - - - - - - - - - - - - - - - -	
Dd	112301 CGAGTTTTGGGTGATGATCGACTCTTCCTCTTTCCAGTCGGATTTTCGATCATGTTAA	112360

Db	Sequence	Position
0Y	765 TGGATATCACTAAATGCGAAAAAGAAATATTTAGAGCACACAAATAAAATTTGGCTCAAG 	824
112361	TGGATTACGGACTAAGGCTACGGAAGTTATTTGGAACANTTCGAACGGATCGC-----	112414

0Y	825	AAAAAAAAATGGAGAAAAATCCACACATATGCTCTGACGCGAGCTCACTTTTATCATTT	---	881
Db	112415	-----AGAGGTTTCTGGAAGTTATCATAAAGATGACGCTCACTTTTCGAGCAACTTTT		1124655

[illegible][illegible]

Db	112526	IGAACTTTGGATCCACCTCAGGACCTACGCCAGTCATTTTGGCTCCGACAAATTCGAAA	1125855
Qy	999	ATTTCGAAAACCTGAAGCATATCTTATTGTGGGAATGATGTCOCAGATGCATCTTGAGA	1058
Db	112586	AAATATCAAAAGCCCTCGATTTTATTATTTACACTCATGAGACAAATCGGACGATGA	1126455
Qy	1059	CAGTTTCCATTCATATCATATAATACATTTATGTGTCCAGTATGGTATATGCGGAAAAGTA	1118
Db	112646	ATCCTTCCCATTCATATTCATATTCATTTTATATGATATCCTGTGCGGTATGCTGGAAGTT	1127055
Qy	1119	TAAATTTGGAATCTGTTCGCC---CAGATGAGTTCACAGGAACATTCAGATGGATGATA	1175
Db	112706	CGGAATCAAGCTGTGGAATCGGCTTGACAGGACACTGAGCCCGGAAATTCAGATGGGAAGGTA	1127655
Qy	1176	CTTGGAGAAAGATCTGCAGAAACCCCTACCGCTTGACTTGTTCACCAATCCCTTGCCA	1235
Db	112766	TTTGAAGGAGAAAGCAGCGGGAATAAATTCCTGATGATGCTTAGCGCGCTTGCCGAGCAA	1128225
Qy	1236	AGAGACATTAGCAAAATTTAAGTAAATCT	1265
Db	112826	GGAGAGAAAGCAATGTTTGAAGTGTCTTTT	112855

RESULT 5	
BC020018	
LOCUS	BC020018
DEFINITION	mus musculus mbt domain containing 1, mRNA (cdna clone MGC:29000)
ACCESSION	BC020018
VERSION	BC020018.1 GI:18043655
KEYWORDS	MGC.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus

REFERENCE
1 (bases 1 to 2878)
Strasberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
AUTHORS

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, I., Max, S.I., Wang, J.J., Hsieh, F., Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Udell, T.B., Toshimori, S., Camiceli, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.W., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huliy, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kelleman, M., Madan, A.C., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shcherbakov, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Gilmwood, J., Schmutz, J., Myers, R.M., Burtfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16699-16903 (2002)
MEDLINE	22388257
PUBMED	12477932
REFERENCE	2 (bases 1 to 2878)
AUTHORS	Strausberg, R.
TITLE	Direct Submission
JOURNAL	Submitted (19-DEC-2001) National Institutes of Health, Mammalian

REMARK	COMMENT
NIH-MGC Project URL: http://mgc.nci.nih.gov	Contact: MGC help desk

Email: cgabbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome


```
misc_feature      /note="assembly_fragment"  
32227. 43573  
misc_feature      /note="assembly_fragment"  
43674. 357577  
misc_feature      /note="assembly_fragment"  
57678. 74056  
misc_feature      /note="assembly_fragment"  
74157. 116911  
misc_feature      /note="assembly_fragment"  
117012. 139343  
misc_feature      /note="assembly_fragment"  
139444. 173294  
misc_feature      /note="assembly_fragment"  
173395. 214213  
misc_feature      /note="assembly_fragment"  
214314. 230185  
misc_feature      /note="assembly_fragment"  
clone_end:T7  
vector_side:right"  
BASE COUNT      70962 a 44539 c 43526 g 69658 t 1500 others  
ORIGIN
```

	Query Match	3.9%;	Score 58.8;	DB 2;	Length 230185;	
	Best Local Similarity	55.3%;	Pred. No. 0.0057;			
	Matches 114;	Conservative	0;	Mismatches 92;	Indels 0;	Gaps 0;
Qy	1279 GTAGACTACGCCCTTGAAGCTGCTGCATCATGTGTGAATAATCAGTTATTATTTGTCACGTACA	1338				
Dd	86438 GTTGGAGATGAAAGATTTGAAAGCTCGTGCAATGAAAAAATCCATTTATGATCTGCTGCACCAC	8649				
Qy	1339 GTGAATACAGTTCATCTGTAAGACATGATAATTCATATTTCAGCGCGCTGGAGAAGAAATTT	1398				
Dd	86498 ATTGAGACCTGCTGTGGAGATCACTTCATATCACCTTTGATVGATGGAGTGAGCAATTT	8655				
Qy	1399 GATGAACATGATGATGTGAGACTCCCATGATATTTCTAACGATAGAGTGGTGAAGCCAC	1458				
Dd	86558 GATTATTGGTGCTGACTATGACTCCCCGAGACATCTTCCAGTAGTGATGCTGCGCTCACA	8661				
Qy	1459 AGTTATGTCTACAACTCCGAAA	1484				
Dd	86618 GGAGATGTATTACAGCCACGAGAA	86643				
RESULT 12						
LOCUS	AX046388					
DEFINITION	Axonin-1, mouse.					
ACCESSION	AX046388					
VERSION	AX046388.1					
KEYWORDS						
SOURCE ORGANISM	Mus musculus (house mouse)					
REFERENCE	Mus musculus Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. Lemischke, I., and Moore, K. Genes that regulate hematopoietic blood forming stem cells and uses thereof Patent: WO 0011168-A 55 02-MAR-2000; Princeton University (US) Location/Qualifiers 1..3590 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090"					
JOURNAL						
FEATURES						
source						
BASE COUNT	1101 a 702 c 797 g 990 t					
ORIGIN						
	Query Match	3.8%;	Score 57.4;	DB 6;	Length 3590;	
	Best Local Similarity	55.2%;	Pred. No. 0.016;			
	Matches 112;	Conservative	0;	Mismatches 91;	Indels 0;	Gaps 0;
y	1276 CGGCTAGAGACTACGCCCTTGAAGCTGCTGCATCATGTGTGAATAATCAGTTATTATTTGTCACACT	1335				

[illegible]

	RESULT	13
	LOCUS	AK000062
	DEFINITION	Homo sapiens cDNA FLJ20055 fis, clone COL00943.
	VERSION	AK000062
	KEYWORDS	AK000062.1 GI:7019904
	SOURCE	cDNA library from testis tissue (full insert sequence).
	ORGANISM	Oligo cloning; Homo sapiens (human)
	REFERENCE	Eukaryote Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo. 1 (sites) Kawabata,A., Hiki,T., Kobayake,N., Inagaki,H., Ikema,Y., Okamoto,S., Ohtani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano.S. NEDO human cDNA sequencing project Unpublished 2 (bases 1 to 2564) Sugano,S., Suzuki,Y., Ota.T., Obayashi.M., Nishi.T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y. Direct Submission
TITLE	JOURNAL	Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Deptment of Virology; Shirokane-dai, 4-6-1, Minato-ku, Tokyo, 108-8639, Japan (E-mail:c dna@ims.u-tokyo.ac.jp, Tel:-81-3-5449-5286, Fax:-81-3-5449-5416)
COMMENT		NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency). Location/Oualifiers 1..2564 /organism="Homo saplens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="COL00943" /tissue_type="colon" /note="cloning vector pME18SF13" 169..1401 /codon_start=1 /protein_id="BAA90919.1" /db_xref="GI:7019905"
CDS		/translation="MGTCWGDISGNCVAVEVPNTDSCLPKEVFMIAGIVKLAVNALLR YEGPNDSGIDPFMCNICGSDIHPCNVACASGRPLPYPRPTOHXYTNMKARVLRTGA KTLPDPDSQKVSESMOIPFKPCAREVEVDKRHLCTRRVAAYESYGRLALYYEEED RITDFGWCHMSPLIHIIGWSRSIGHRKSDITRKOGDHFTDPHFPAVKVEDQSSE WFPEGMLEAIDPLINTICVATIRLVADGLFIIIGDSEAAGSDMFCYAHTSPSI EPVGCEIMILEETLPPRGYTILPKPWFDYLFREGSIAAPVLAENKDVPNHGFRGMFL EAVIDMPRELCAVTATRYRIITHLLRHIFDGWEERYDDQWDCCESPDIYPVMCQLTGYQ LOFPASCKLYVRGVLL"
BASE COUNT		832 A 482 C 560 G 690 T
ORIGIN		

	Query Match	3.7%:	Score 55.8:	DB 9:	Length 2564:
	Best Local Similarity 54.7%:	Pred.	No. 0.038:		
	Matches 111:	Conservative 0:	Mismatches 92:	Indels 0:	Gaps 0:
QY	1276	CGGGTAGGACATACCCCTTGAAGCTGCACATGTGTGAAATCAGTTATTTGTCCAGCT	1335		
Db	1153	CGTGTAGGAATGAAATTAGGAAGCAGTGTGATCTCATGGAGCCACGTTTAAATGTGTGAGCC	1212		
QY	1336	ACAATGAAATCAGTTTCATGGAAGCTGATTAATGTCAATTTTCGACGCTGGGATGGAAGAA	1395		
Db	1213	ACAATTAATCGAATTTATTCATGTCCTCTTGAGATACATTTTGTATGGATGGGAAGAAGAG	1272		
QY	1396	TTTGTAAACATGTATGTATGTGAGACTTCCCATATATTTACCGATAGAGATGGTGTGAAGCG	1455		
Db	1273	TATGATCAAGTGGGTAGACTGTGATTCACACTGATCTATCTGTAGAGGTGGTGTCAGTTA	1332		
QY	1456	CACAGTTATGTCTACAACCTCC	1478		
Db	1333	ACTGGATATCAACTACAGCCTCC	1355		

RESULT	14
AC005839/c	
LOCUS	AC005839 128328 bp DNA linear PRI 30-DEC-1998
DEFINITION	Homo sapiens chromosome 17, clone hPK_481_C_4, complete sequence.
ACCESSION	AC005839
VERSION	AC005839.1 GI:4079626
KEYWORDS	HTG.
SOURCE	Homo sapiens
ORGANISM	Homo sapiens (human)

REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS
1 (bases 1 to 128328)	Birtlen,B.,	Linton,L.,	Nusbaum,C.	and Lander,E.	
	Homo sapiens chromosome 17, clone hRPK_481_C_4				
	Unpublished				
2 (bases 1 to 128328)	Birtlen,B.,	Linton,L.,	Nusbaum,C.,	Lander,E.,	Allen,N.,
	Anderson,M.,				

TITLE Direct Submission
JOURNAL Submitted (21-OCT-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 128528)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,

TITLE Direct Submission
JOURNAL Submitted (30-DEC-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT On Dec 30, 1998 this sequence version replaced gi:4049329. All repeats were identified using RepeatMasker: Smtl, A.F.A. & Green, P. (1996-1997) <http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Only the last 128328 base pairs of this clone are being submitted The remainder overlaps with accession number AC005920 (WICGR project L444).

```
FEATURES
source
location/Qualifiers
1..128328
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="hRPK_481_C_4"
/clone_1b="RPCL-11 human BAC library"
repeat_region
304..437
/rpt_family="Alusg"
repeat_region
438..725
/rpt_family="Aluy"
repeat_region
728..757
/rpt_family="(CAAA)n"
repeat_region
758..906
/rpt_family="Alusg"
repeat_region
1398..1707
/rpt_family="Alusx"
repeat_region
2026..2323
/rpt_family="AluJo"
repeat_region
2327..2607
/rpt_family="Aluy"
repeat_region
2623..2905
/rpt_family="AluJb"
repeat_region
2937..3243
/rpt_family="Alusx"
repeat_region
3363..3454
/rpt_family="FLAM_C"
repeat_region
complement(3455..3468)
/rpt_family="(TAGA)n"
repeat_region
3469..3588
/rpt_family="(CTAG)n"
repeat_region
complement(3589..3619)
/rpt_family="(TAGA)n"
repeat_region
complement(3620..3904)
/rpt_family="Alusp"
repeat_region
complement(3905..3917)
/rpt_family="(TAGA)n"
repeat_region
complement(4137..4436)
/rpt_family="Alusg"
repeat_region
complement(4446..4576)
/rpt_family="AluJo"
repeat_region
5175..5222
/rpt_family="AT_rich"
repeat_region
complement(5715..6021)
/rpt_family="Alusx"
repeat_region
6643..7083
/rpt_family="L2"
repeat_region
7084..7395
/rpt_family="Alusg"
repeat_region
7396..7735
/rpt_family="L2"
repeat_region
7736..7941
/rpt_family="Alusg/x"
repeat_region
7945..8245
/rpt_family="Alusp"
repeat_region
8246..8577
/rpt_family="L2"
repeat_region
8578..8887
/rpt_family="Alusx"
repeat_region
8888..9118
/rpt_family="L2"
repeat_region
complement(9708..10019)
/rpt_family="Alusx"
```

repeat_region	complement(10027. .10066) /rpt_family="(TAA)n"
repeat_region	complement(10067. .10196) /rpt_family="FRAM"
repeat_region	complement(10563. .10612) /rpt_family="L2"
repeat_region	complement(10879. .11191) /rpt_family="AluSg"
repeat_region	complement(11324. .11630) /rpt_family="AluSg"
repeat_region	complement(12443. .12762) /rpt_family="AluSx"
repeat_region	12842. .13141 /rpt_family="AluSg"
repeat_region	13147. .13434 /rpt_family="AluSg"
repeat_region	complement(14208. .14527) /rpt_family="AluB"
repeat_region	14998. .15291 /rpt_family="AluY"
repeat_region	15300. .15575 /rpt_family="AluSc"
unsure	15527. .15536 /note="Single-stranded terminator coverage."
repeat_region	15576. .15596 /rpt_family="(CAA)n"
repeat_region	15597. .15615 /rpt_family="AluSc"
repeat_region	15625. .15606 /rpt_family="AluSg"
repeat_region	15907. .15934 /rpt_family="(CAA)n"
repeat_region	15944. .16033 /rpt_family="L1P82"
repeat_region	16036. .16153 /rpt_family="(TA)n"
repeat_region	complement(16165. .16459) /rpt_family="AluSx"
repeat_region	complement(16536. .16613) /rpt_family="L2"
repeat_region	16723. .16997 /rpt_family="AluSg"
repeat_region	17461. .17779 /rpt_family="AluB"
repeat_region	complement(17976. .17995) /rpt_family="(CAA)n"
repeat_region	18466. .18489 /rpt_family="AT-rich"
repeat_region	18652. .19114 /rpt_family="L1MC5"
repeat_region	19115. .19428 /rpt_family="AluB"
repeat_region	19429. .19565 /rpt_family="L1MC5"
repeat_region	19576. .19708 /rpt_family="FLAM_C"
repeat_region	19777. .19903 /rpt_family="L1ME"
repeat_region	19904. .20203 /rpt_family="AluY"
repeat_region	20204. .20222 /rpt_family="L1ME"
repeat_region	20677. .20802 /rpt_family="HALL1"
repeat_region	20996. .21154 /rpt_family="L1MC5"
repeat_region	complement(21185. .21272) /rpt_family="L1MC5"
repeat_region	21497. .21805 /rpt_family="AluY"
repeat_region	21811. .21923 /rpt_family="AluSg"
repeat_region	21950. .22149

	repeat_region	/rpf_family="Alusg"
		22150..22265
		/rpf_family="(TA)n"
	repeat_region	22429..22554
Query Match	3.7%;	Score 55.8; DB 9; Length 128328;
Best Local Similarity	54.7%;	Pred. No. 0.029;
Matches 111; Conservative	0;	Mismatches 92; Indels 0; Gaps 0;
Oy	1276	CGGCTAGACATCAGCCTTGAAGCTCTGACATGTGTGAAATACAGTTTATTGTCAGCT 1335
Db	119478	CGTGATGAATGAATTAAGAACCAATGATCTCATAGAGCCACGTTAATATGTGTAGCC 1194119
Oy	1336	ACAGTGAATCAGTCTCATGAGAAGACTGATTAATGTCAATTTGACGCGCTGGAGTGAAGAA 1395
Db	119418	ACAGTAACCTCGAATTATTCATCAGTCGTCTCTTGAGAGTACATTTTGGATGAGTGGGAAGAAG 119359
Oy	1396	TTTGTATGAACGTATGATGATGTGAGACTCCCATGATATTCTACCGATAGATGGTGTGAAGG 1455
Db	119358	TATGATTCAGTGGGTGAGACTGTGTAGACACTGACCTATACCTGTAGGGTGTGTGCTGTTA 119299
Oy	1456	CACAGTTATGTCTCAACCTCC 1478
Db	119298	ACTGATATCAACTACAGCCTCC 119276

```

LOCUS       199301 bp      DNA      linear      HTG 15-JAN-2003
DEFINITION  Danio rerio clone DKER-27E7, *** SEQUENCING IN PROGRESS ***, 37
unordered pieces.
ACCESSION   BX088589
VERSION     BX088589.1  GI:27764200
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Danio rerio (zebrafish)
ORGANISM    Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 199301)
Burton,J.
Direct Submission
Submitted (14-JAN-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish.help@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zK27E7
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 182613 bases at least Q40
Consensus quality: 187913 bases at least Q30
Consensus quality: 191719 bases at least Q20
Insert size: 195701; sum-of-contrigs
Insert size: 237259; 8.1% error; agarose-fp
Quality coverage: 2.54x in Q20 bases; sum-of-contrigs quality
coverage: 2.29x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 37 contrigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contrigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
*       7923: contrig of 7923 bp in length
*       1
*       7924      8023: gap of 100 bp

```

FEATURES	source
* 8024	10536: contig of 2513 bp in length
* 10537	10636: gap of 100 bp
* 10637	12687: contig of 2051 bp in length
* 12688	12787: gap of 100 bp
* 12788	13386: contig of 2599 bp in length
* 15386	15486: gap of 100 bp
* 15487	21116: contig of 5630 bp in length
* 21117	21216: gap of 100 bp
* 21217	27096: contig of 5880 bp in length
* 27097	27196: gap of 100 bp
* 27197	31488: contig of 4292 bp in length
* 31489	31588: gap of 100 bp
* 31589	35763: contig of 4175 bp in length
* 35764	35863: gap of 100 bp
* 35864	37963: contig of 2100 bp in length
* 37964	38063: gap of 100 bp
* 38064	44724: contig of 6661 bp in length
* 44725	44824: gap of 100 bp
* 44825	49012: contig of 4188 bp in length
* 49013	49112: gap of 100 bp
* 49113	56246: contig of 7134 bp in length
* 56247	56346: gap of 100 bp
* 56347	58741: contig of 2395 bp in length
* 58742	58841: gap of 100 bp
* 58842	62628: contig of 3787 bp in length
* 62629	62728: gap of 100 bp
* 62729	73857: contig of 11129 bp in length
* 73858	73957: gap of 100 bp
* 73958	76876: contig of 2919 bp in length
* 76877	76976: gap of 100 bp
* 76977	83827: contig of 6851 bp in length
* 83828	83927: gap of 100 bp
* 83929	92033: contig of 8106 bp in length
* 92034	92133: gap of 100 bp
* 92134	95917: contig of 3784 bp in length
* 95918	96017: gap of 100 bp
* 96018	104147: contig of 8130 bp in length
* 104148	104247: gap of 100 bp
* 104248	108202: contig of 3955 bp in length
* 108203	108302: gap of 100 bp
* 108303	130167: contig of 21865 bp in length
* 130168	130267: gap of 100 bp
* 130268	132754: contig of 2487 bp in length
* 132755	132854: gap of 100 bp
* 132855	136967: contig of 4113 bp in length
* 136968	137067: gap of 100 bp
* 137068	146448: contig of 9381 bp in length
* 146449	146548: gap of 100 bp
* 146549	150886: contig of 4338 bp in length
* 150887	150986: gap of 100 bp
* 150987	153567: contig of 2581 bp in length
* 153568	153667: gap of 100 bp
* 153668	155805: contig of 2138 bp in length
* 155806	155905: gap of 100 bp
* 155906	158786: contig of 2881 bp in length
* 158787	158886: gap of 100 bp
* 158887	166773: contig of 7887 bp in length
* 166774	166873: gap of 100 bp
* 166874	175651: contig of 8778 bp in length
* 175652	175751: gap of 100 bp
* 175752	179746: contig of 3995 bp in length
* 179747	179846: gap of 100 bp
* 179847	182392: contig of 2546 bp in length
* 182393	182492: gap of 100 bp
* 182493	186389: contig of 3897 bp in length
* 186390	186489: gap of 100 bp
* 186490	188775: contig of 2286 bp in length
* 188776	188875: gap of 100 bp
* 188876	194469: contig of 5594 bp in length
* 194470	194569: gap of 100 bp
* 194570	199301: contig of 4732 bp in length.
Location/Qualifiers	
1. .199301	

misc_feature	/organism="Danio rerio"
	/mol_type="genomic DNA"
	/db_xref="taxon:7955"
	/clone="DKFZ-27E7"
	/clone_lib="DanioKey"
	1. .7923
misc_feature	/note="assembly_fragment:00057
	fragment_chain:1"
	8024. 10536
misc_feature	/note="assembly_fragment:00635
	fragment_chain:1"
	10637. 12687
misc_feature	/note="assembly_fragment:00201
	fragment_chain:1"
	12788. 15386
misc_feature	/note="assembly_fragment:00309
	fragment_chain:1"
	15487. 21116
misc_feature	/note="assembly_fragment:00082
	fragment_chain:2"
	21217. 27096
misc_feature	/note="assembly_fragment:00610
	fragment_chain:2"
	27197. 31488
misc_feature	/note="assembly_fragment:00367
	fragment_chain:2"
	31589. 35763
misc_feature	/note="assembly_fragment:00292
	fragment_chain:3"
	35864. 37963
misc_feature	/note="assembly_fragment:00675
	fragment_chain:3"
	38064. 44724
misc_feature	/note="assembly_fragment:00404
	fragment_chain:3"
	44825. 49012
misc_feature	/note="assembly_fragment:00317
	fragment_chain:4"
	49113. 56246
misc_feature	/note="assembly_fragment:00033
	fragment_chain:4"
	56347. 58741
misc_feature	/note="assembly_fragment:00302
	fragment_chain:4"
	58842. 62628
misc_feature	/note="assembly_fragment:00369
	fragment_chain:5"
	62729. 73857
misc_feature	/note="assembly_fragment:00424
	fragment_chain:5"
	73958. 76876
misc_feature	/note="assembly_fragment:00469
	fragment_chain:5"
	76977. 83827
misc_feature	/note="assembly_fragment:00581
	fragment_chain:6"
	83928. 92033
misc_feature	/note="assembly_fragment:00518
	fragment_chain:6"
	92134. 95917
misc_feature	/note="assembly_fragment:00059
	fragment_chain:6"
	96018. 104147
misc_feature	/note="assembly_fragment:00137
	fragment_chain:7"
	104248. 108202
misc_feature	/note="assembly_fragment:00510
	fragment_chain:7"
	108303. 130167
misc_feature	/note="assembly_fragment:00213
	fragment_chain:8"
	130268. 132754
misc_feature	/note="assembly_fragment:00313

